IMPERIAL INSTITUTE
OF
AGRICULTURAL RESEARCH PUSA

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND
PHYSICAL CHARACTER.

VOL. CXIX.

LONDON:
PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE,
PRINTERS IN ORDINARY TO HIS MAJESTY.
JULY, 1928.

LONDON:
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. CXIX.

Minutes of Meetings of May 10, June 7, 14, 21, 28, 1928.

No. A 781.—May 1, 1928.

PAGE

The Saha Theory and the Conductivity of Flames containing Alkali Metal Vapours. By H. A. Wilson, F.R.S. 1

The Light of the Night Sky: its Intensity Variations when Analysed by Colour Filter.—III. By Lord Rayleigh, F.R.S. 11

On the Brownian Displacements and Thermal Diffusion of Grains suspended in a Non-Uniform Fluid. By S. Chapman, F.R.S. 34

The Molecular Displacements in Diffusing Gas-Mixtures. By S. Chapman, F.R.S. 55

The Excitation of Soft X-Rays.—II. By O. W. Richardson, F.R.S., and F. C. Chalklin 60

Intensity Measurements in the Secondary Spectrum of Hydrogen. By L. S. Ornstein, W. Kapucinski and Miss J. G. Eymers. Communicated by O. W. Richardson, F.R.S. 83

An Experimental Study of the Flow of Water in Pipes of Rectangular Section. By S. J. Davies and C. M. White. Communicated by E. V. Appleton, F.R.S. 92

On Quantitative Measurements in Frictional Electricity. By W. A. Macky. Communicated by R. Whiddington, F.R.S. 107

The Structure of Cyanite, Al_2SiO_5. By W. H. Taylor and W. W. Jackson. Communicated by W. L. Bragg, F.R.S. 132

The Influence of Vortices upon the Resistance experienced by Solids moving through a Liquid. By W. G. Bickley. Communicated by G. I. Taylor, F.R.S. 146

The Origin of Speech—A Hypothesis. By Sir Richard A. S. Paget, Bart. Communicated by Sir William Bragg, F.R.S. 157

Electron Emission in Intense Electric Fields. By R. H. Fowler, F.R.S., and L. Nordheim 173

No. A 782.—June 1, 1928.

On Line-Integrals of the Diurnal Magnetic Variations. By A. T. Price and S. Chapman, F.R.S. 182

A New Interference Method of Measuring the Surface Area of Film Catalysts. Part I.—The Theory. By F. H. Constable. Communicated by T. M. Lowry, F.R.S. 196

PAGE

A New Interference Method of Measuring the Surface Area of Film Catalysts. Part II.—Nickel. A Method of Preparation of the Film, and an Apparatus for Activation, and Study of the Surface Area. By F. H. Constable. Communicated by T. M. Lowry, F.R.S. 202

The Nodal Cubic Surfaces and the Surfaces from which they are derived by Projection. By G. Timms. Communicated by H. F. Baker, F.R.S. 213

The Spark Spectrum of Copper. By A. C. Menzies. Communicated by Prof. R. Whiddington, F.R.S. 249

The Existence of Intermetallic Compounds in the Vapour State. The Spectra of the Alkali Metals, and of their Alloys with each other. By J. M. Walter and S. Barratt. Communicated by T. R. Merton, F.R.S. (Plate 1) 257

The Theory of Rayleigh's Principle as applied to Continuous Systems. By G. Temple. Communicated by S. Chapman, F.R.S. 276

On Phenomena occurring in the Condensation of Molecular Streams on Surfaces. By J. D. Cockcroft. Communicated by Sir Ernest Rutherford, P.R.S. (Plates 2 and 3) 293

The Stark Effect of the Fine-Structure of Hydrogen. By R. Schlapp. Communicated by C. G. Darwin, F.R.S. 313

Statistical Experiments on the Motion of Electrons in Gases. By R. d'E. Atkinson. Communicated by F. R. Lindemann, F.R.S. 335

Observations on the Band Spectra of Mercury. By Lord Rayleigh, F.R.S. (Plate 4) 349

The Study of the Specific Resistance of Bismuth Crystals and its Change in Strong Magnetic Fields and some Allied Problems. Parts I—III. By P. Kapitza. Communicated by Sir Ernest Rutherford, P.R.S. (Plates 5–8) 358

The Polarisation of Radio Waves. By J. Hollingworth. Communicated by permission of the Radio Research Board, by Sir Joseph Petavel, F.R.S. 444

Gaseous Combustion at High Pressures. Part X.—The Co-volume Corrections, Maximum Temperatures and Dissociation of Steam and Carbon Dioxide in Explosions. By D. M. Newitt. Communicated by Prof. W. A. Bone, F.R.S. 464

No. A 783.—July 2, 1928.

Erosion by Water-hammer. By S. S. Cook, F.R.S. (Plate 9) 481

Solid Dipleidoscope Prisms. By C. V. Boys, F.R.S. 489

The Mechanism of Chemical Change. I.—Promotion and Arrest of the Mutarotation of Tetra-Acetylglucose in Ethyl Acetate. By T. M. Lowry, F.R.S., and G. Owen 505

Note on Total Reflexion of Electric Waves at the Interface between Two Media. By H. M. Macdonald, F.R.S. 523

PAGE

A Critical-Absorption Photometer for the Study of the Compton Effect. By C. V. Raman, F.R.S., and C. M. Sogani. (Plate 10) 526

The Emission of Secondary Electrons and the Excitation of Soft X-Rays. By O. W. Richardson, F.R.S 531

The Photo-Electric Properties of some Metals in the Soft X-Ray Region. By L. P. Davies. Communicated by O. W. Richardson, F.R.S 543

Condensation Phenomena at Different Temperatures. By C. F. Powell. Communicated by C. T. R. Wilson, F.R.S 553

The Effect of Temperature on the Viscosity of Neon. By R. S. Edwards. Communicated by H. L. Callendar, F.R.S 578

The Homogeneous Reaction between Hydrogen and Oxygen. By C. H. Gibson and C. N. Hinshelwood. Communicated by Sir Harold Hartley, F.R.S 591

On Series in the Spark Spectra of Germanium. By K. R. Rao and A. L. Narayan. Communicated by Lord Rayleigh, F.R.S. (Plate 11) 607

The Structure of Thin Films. Part XI.—Oxygenated Derivatives of Benzene. By N. K. Adam. Communicated by Sir William Hardy, F.R.S 628

The Dependance of the Photographic Action of β-Rays on their Velocity. By C. D. Ellis and G. H. Aston. Communicated by Sir Ernest Rutherford, P.R.S 645

Experiments on the Diffraction of Cathode Rays.—II. By G. P. Thomson. Communicated by Sir Joseph Thomson, F.R.S. (Plate 12A) 651

The Diffraction of Cathode Rays by Thin Celluloid Films. By A. Reid. Communicated by Sir Joseph Thomson, F.R.S. (Plate 12B) 663

The Diffraction of Cathode Rays by Thin Films of Copper, Silver and Tin. By R. Ironside. Communicated by Sir Joseph Thomson, F.R.S. 668

The Scattering of Cathode Rays. By B. F. J. Schonland. Communicated by Sir Ernest Rutherford, P.R.S 673

Metal Overvoltage Measurements with the Cathode Ray Oscillograph. By E. Newbery. Communicated by Sir Ernest Rutherford, P.R.S. (Plate 13) 680

A Revision of the Theory of Transfer Resistance. By E. Newbery. Communicated by Sir Ernest Rutherford, P.R.S 686

On the Kinetic Method in the New Statistics and its Application in the Electron Theory of Conductivity. By L. W. Nordheim. Communicated by R. H. Fowler, F.R.S. 689

The Fluorescence of Mercury Vapour. By R. W. Wood, For. Mem. R.S., and V. Voss. (Plate 14) 698

An Improved Method of Ultra-Violet Polarimetry. Anomalous Rotary Dispersion of Sodium Tartrate. By T. M. Lowry, F.R.S., and M. A. Vernon 706

OBITUARY NOTICES.

PAGE
Alfred George Greenhill (with portrait) i
Gösta Mittag-Leffler (with portrait) v
Svante August Arrhenius (with portrait) ix
Paul Heinrich Ritter von Groth (with portrait) xx
William Henry Dines (with portrait) xxiii

Index xxxiii

PROCEEDINGS OF
THE ROYAL SOCIETY.

Section A.—Mathematical and Physical Sciences.

The Saha Theory and the Conductivity of Flames containing Alkali Metal Vapours.

By H. A. Wilson, F.R.S., Rice Institute, Houston, Texas.

It was shown by A. A. Noyes and the writer* that the equilibrium constants for the reactions $M = M_+ + e$, where M denotes an atom of an alkali metal and e an electron, can be deduced from measurements of the electrical conductivity of flames, at 2000° K., containing the metal, and that the values so obtained agree approximately with those calculated by means of Saha's well-known theory.†

This conclusion has since been fully confirmed by B. T. Barnes‡ and by J. A. J. Bennett.§ Bennett found that the equilibrium constants deduced from his measurements agreed with Saha's theory at temperatures near 2000° K., but that at lower temperatures the theoretical values were smaller than those given by the experiments.

These determinations of the equilibrium constants have been criticised by E. Marx in the new edition of his book 'Flammenleitung' (Leipzig, 1927).

Marx points out that two assumptions were made in the calculations of the equilibrium constants which may not be correct. It was assumed that the alkali metal salt sprayed into the flame is completely converted into metal vapour and it was assumed that the negative carriers in flames are all electrons.

Let us then suppose that only a fraction f of the salt introduced into the flames is reduced to the metallic state and that only a fraction q of the carriers of the negative electricity are free electrons at any instant.

Consider the reaction $M = M_+ + e$. Let p_1 denote the partial pressure of the neutral metal atoms in the flame, p_3 that of the electrons and p_2 that of the positively charged metal atoms. If K is the equilibrium constant for this reaction, we have $\mathrm{K} = p_2p_3/p_1$. Also suppose some of the molecules of the flame gases are dissociated into positive ions and electrons thus, $\mathrm{F} = \mathrm{F}_+ + e$, and let $\mathrm{K}' = p_2'p_3/p_1'$, where K′ is the equilibrium constant and $p_1'p_2'$ the partial pressures of F and F_+.

---

* 'J. Amer. Chem. Soc.,' vol. 44, p. 2806 (1922).

† Saha, 'Phil. Mag.,' vol. 40, pp. 478, 809 (1920); vol. 41, p. 267 (1921).

‡ 'Phys. Rev.,' vol. 23, p. 189 (1924).

§ 'Phil. Mag.,' vol. 3, p. 127 (1927).

The conductivity of the flame is equal to $e(n_1k_1 + n_2k_2 + n_3k_3)$, where e is the protonic charge, n_1 n_2 n_3 the numbers per cubic centimetre of the positive ions, the negative ions and the electrons, and k_1 k_2 k_3 the respective mobilities of the ions and electrons. We call the negatively charged atoms or molecules the negative ions to distinguish them from the free electrons.

The fraction q is equal to $n_3/(n_2 + n_3)$ and we have also $n_1 = n_2 + n_3$, so that the conductivity is equal to $n_1e(k_1 + (1-q)k_2 + qk_3)$. Now k_3 is very much larger than k_1 and k_2, so that the terms k_1 and $(1-q)k_2$ may be neglected unless q is a very small fraction. The average mobility of all the ions $\bar{k}$ is given by $\bar{k} = (n_1k_1 + n_2k_2 + n_3k_3)/(n_1 + n_2 + n_3)$ or $\bar{k} = \frac{1}{2}(k_1 + (1-q)k_2 + qk_3)$. This equation gives $q = (2\bar{k} - k_1 - k_2)/(k_3 - k_2)$. The ratio of qk_3 to $k_1 + k_2$ is therefore given by

$$qk_3/(k_1 + k_2) = (2\bar{k} - k_1 - k_2)\,k_3/(k_3 - k_2)\,(k_1 + k_2).$$

Provided k_1 and k_2 are small compared with both k_3 and $\bar{k}$, this reduces to $qk_3/(k_1 + k_2) = 2\bar{k}/(k_1 + k_2)$. Now k_1 and k_2 are believed to be from 1 to 20 cms./sec. per volt/cm. and $\bar{k}$ to be not less than about 500, so that $qk_3/(k_1 + k_2) = 500$ to 25. Thus qk_3 is large compared to k_1 and k_2, so that the conductivity is nearly equal to n_1eqk_3 or en_3k_3. Also $q = 2\bar{k}/k_3$. According to Marx $\bar{k}$ is about 500, so that if k_3 is 2500 then $q = 0{\cdot}4$ and if $k_3 = 25{,}000$ then $q = 0{\cdot}04$. There are, however, good reasons for believing that $q = 1$ in flames at about 2000° K., in which case $\bar{k} = k_3/2$. Since $\bar{k}$ is the average mobility of all the carriers, the average mobility of the negative carriers is nearly $2\bar{k}$ because k_1/k_3 is negligible.

The partial pressure of the electrons is equal to $p_3 = n_3\mathrm{R}_1\mathrm{T}$, where R_1 is the gas constant for one molecule and T the absolute temperature. Hence, if c denotes the conductivity, we have $p = c\mathrm{R}_1\mathrm{T}/k_3e$, or if N denotes the number of molecules in one mol, so that $\mathrm{NR}_1 = \mathrm{R}$, the gas constant for one mol, then $p_3 = c\mathrm{RT}/k_3\mathrm{F}$, where $\mathrm{F} = \mathrm{N}e$ denotes the Faraday, or if $\mathrm{A} = \mathrm{RT}/k_3\mathrm{F}$, then $p_3 = \mathrm{A}c$.

If the solution sprayed into the flame contains g grams of salt per litre, then we may write $\mathrm{B}fg = p_1 + p_2$, where B is a constant and f the fraction of the metal in the salt which is reduced to the metallic state in the flame. We have also $p_3 = q(p_2 + p_2')$, since the electrons and negative ions together are equal to the positive ions.

The equations obtained, on eliminating p_1, p_2, p_3 and p_2', give

$$\frac{\mathrm{A}c}{q} = \frac{\mathrm{KB}fg}{\mathrm{K} + \mathrm{A}c} + \frac{\mathrm{K}'p_1'}{\mathrm{A}c}.$$

Now let $c = c_0$ and $q = q_0$ when $g = 0$, and then, provided p_2'/p_1' is a very small fraction so that p_1' may be regarded as a constant, we have $\mathrm{A}c_0/q_0 = \mathrm{K}'p_1'/\mathrm{A}c_0$. Then putting $x = c/c_0$ we get

$$\frac{gx}{x^2 - q/q_0} = \frac{\mathrm{A}c_0}{\mathrm{B}fg} + \frac{\mathrm{A}^2c_0^2x}{\mathrm{KB}fq}.$$

If we put $f = q = 1$, this becomes

$$\frac{gx}{x^2 - 1} = \frac{\mathrm{A}c_0}{\mathrm{B}} + \frac{\mathrm{A}^2c_0^2x}{\mathrm{KB}},$$

which is the equation obtained by the writer* and used by A. A. Noyes and the writer, B. T. Barnes and J. A. J. Bennett.

When g is zero or very small, so that x is small, we have $q = q_0$, and when x is large q/q_0 is negligible compared with x^2, so that q/q_0 may be replaced by unity without appreciable error in any case. It is probable, moreover, that q is practically equal to q_0 for any value of g. We have therefore

$$\frac{gx}{x^2 - 1} = \frac{\mathrm{A}c_0}{\mathrm{B}fq} + \frac{\mathrm{A}^2c_0^2}{\mathrm{KB}fq}x.$$

It was shown by the writer (*loc. cit.*) that the experimentally found values of x for different values of g approximately satisfy an equation of the form

$$\frac{gx}{x^2 - 1} = b + ax,$$

where b and a are constants. The results obtained by B. T. Barnes and J. A. J. Bennett are consistent with this result. A. B. Bryan,† who made some measurements of x and g with an apparatus practically identical with that used by the writer, found the relation between $gx/(x^2 - 1)$ and x to be nearly but not quite linear with cæsium and potassium salts. The following table contains the writer's and also A. B. Bryan's results for cæsium chloride. Bryan's values of

---

* 'Phil. Trans.,' A, vol. 216, p. 63 (1915).

† 'Phys. Rev.,' vol. 17, p. 54 (1921).

g have been multiplied by 0·73 to allow for the difference between the sprayers. The calculated values of g are those given by the equation $gx/(x^2 - 1) = 10 + x$.

x.	g.	g (cal.).	Percentage difference.
1	0	0	
2·72‡	0·00292	0·0030	+ 2·7
2·88	0·0032	0·0033	+ 2·0
5·72	0·008	0·0087	+ 8·1
8·90	0·016	0·0166	+ 3·8
12·25‡	0·0292	0·0270	+ 7·5
13·5	0·032	0·0315	– 1·6
22·7	0·080	0·074	– 7·5
32·8	0·16	0·14	–12·5
55·0‡	0·292	0·358	+23·0
85·2	0·80	0·81	+ 1·3
155‡	2·92	2·56	–12·3
282	8·0	8·2	+ 2·5
883	80	79	– 1·3

Bryan's results are indicated by the sign ‡. With one exception they agree with the theory about as well as those of the writer. Bryan's results for sodium do not agree with the equation $gx/(x^2 - 1) = b + ax$. In the case of sodium b is negligible so that $g = ax^2$ very nearly unless x is very small. It has been shown by a number of different observers that the conductivity due to sodium salts varies nearly as the square root of the concentration, so that it seems probable that Bryan's results for sodium were affected by some unsuspected source of error. His conductivities for dilute sodium solutions are much larger than those obtained by other observers, so that very possibly his dilute sodium solutions were accidentally contaminated with traces of potassium or cæsium. The following table gives Arrhenius'* and the writer's results for sodium salts.

Arrhenius (1891).			H. A. Wilson (1915).		
Mols per litre. n.	Current. i.	$i/\sqrt{n}$.	g.	x.	$x/\sqrt{g}$.
1/256	2·0	32	0·032	1·73	9·67
1/64	4·2	33·6	0·16	2·95	7·37
1/16	8·0	32	0·8	6·06	6·78
1/4	17·7	35·4	4·0	13·7	6·85
1	40·3	40·3	20·0	30·3	6·78

---

* 'Wied. Ann.,' vol. 42, p. 18 (1891).

These results show that the conductivity varies nearly as $\sqrt{n}$, and this result has been confirmed by Moreau and others.

Assuming that the conductivity varies with the concentration in accordance with the equation $gx/(x^2 - 1) = b + ax$ at any rate when g is not very large, we have $b = \mathrm{A}c_0/\mathrm{B}fq$ and $a = \mathrm{A}^2c_0{}^2/\mathrm{KB}fq$. Eliminating $\mathrm{B}fq$ we get $\mathrm{K} = c_0\mathrm{A}b/a$ or $\mathrm{K} = c_0\mathrm{RT}b/k_3\mathrm{F}a$. This is the equation used by A. A. Noyes and the writer, so that it appears that the assumptions $f = 1$ and $q = 1$ to which Marx objects do not make any difference to this equation. Thus this expression for the equilibrium constant in terms of quantities which can be measured is independent of the fraction of the salt reduced to metal in the flame and of the fraction of the negative ions which are free electrons in the flame. The calculation of the equilibrium constants by A. A. Noyes and the writer is therefore unaffected even if the fractions f and q are not equal to one as they assumed.

Barnes and Bennett both determined K in another way. If we eliminate $\mathrm{A}c_0$ instead of $\mathrm{B}fq$ we get $\mathrm{K} = fq\mathrm{B}b^2/a$. Barnes and Bennett used this equation with $fq = 1$. They determined the constants B, b, and a and calculated K by $\mathrm{K} = \mathrm{B}b^2/a$. They both got values of K agreeing with Saha's theory at about 2000° K., which suggests that at this temperature qf is nearly equal to unity for cæsium and rubidium, the metals which they used. Bennett found that at lower temperatures his values of K were greater than those given by Saha's theory. For example, at 1570° K. he found $\mathrm{K} = 6 \cdot 9 \times 10^{-12}$, whereas the Saha theory gives $1 \cdot 1 \times 10^{-12}$ at this temperature. This may well be due to the factor qf becoming less than unity at lower temperatures. We should expect the fraction of the salt reduced to metal to diminish with falling temperature and also the fraction of the negative carriers which are electrons to diminish.

Bennett's results on rubidium enable the equilibrium constant at 1920° K. to be calculated by means of the equation $\mathrm{K} = c_0\mathrm{RT}b/k_3\mathrm{F}a$, since he determined c_0, b and a. At 1920° K. he got $c_0 = 3 \cdot 19 \times 10^{-7}$ reciprocal ohms and $b/a = 2 \cdot 84$. This may be compared with the value $b/a = 2 \cdot 25$ found by the writer at a temperature estimated to be 2000° K.

Taking for k_3 the value 2600 cms./sec. per volt/cm. given by Watt's* measurements of the Hall Effect in flames with small electric fields, we get

$$\mathrm{K} = 3 \cdot 19 \times 10^{-7} \times 82 \cdot 07 \times 1920 \times 2 \cdot 84/2600 \times 96500 = 5 \cdot 7 \times 10^{-10}.$$

Saha's theoretical expression for K gives $5 \cdot 5 \times 10^{-10}$ at 1920° K. Thus it appears that Bennett's results at 1920° K. agree approximately with Saha's theory when the expression for K used by A. A. Noyes and the writer, which is independent of the assumptions $f = 1$ and $q = 1$, is used.

---

* 'Phys. Rev.,' vol. 25, p. 73 (1925).

The equations $K = fqBb^2/a$ and $K = c_0RTb/k_3Fa$ give $k_3 = c_0RT/FBbfq$. Bennett and Barnes used this equation, with $fq = 1$ to calculate the mobility of the electrons from the values of c_0, B and b which they obtained. Barnes got $k_3 = 8000$ at about 2000° K. and Bennett got $k_3 = 2700$ at 1920° K. Bennett got values of k_3 at lower temperatures, but since his values of K at lower temperatures do not agree with those calculated on Saha's theory, which may be due to the factor fq being less than unity, his values of k_3 at lower temperatures cannot be regarded as of much value. If we assume that Saha's theory gives the correct value of K at any temperature, we can calculate the values of the factor fq from Bennett's results. The following table gives the results obtained in this way:—

T.	Log_{10} K. (Saha)	Log_{10} K. (Bennett)	fq.
1920	− 9·26	− 9·26	1·0
1870	− 9·58	− 9·51	0·93
1720	−10·68	−10·35	0·47
1670	−11·05	−10·76	0·54
1570	−11·94	−11·16	0·16

Not much weight can be given to the above values of fq, for it is quite possible that the differences between Bennett's values of K at the lower temperatures and those given by Saha's theory are mainly due to experimental errors and that fq is really nearly equal to unity at any temperature above 1500° K. This question cannot be decided until more accurate measurements become possible.

It is important to note that the mobilities deduced by Barnes and Bennett are the mobilities of the free electrons. The conductivity is equal to $(n_1 + n_2 + n_3)e\bar{k}$, where $\bar{k}$ is the average mobility of all the ions including the free electrons, but this is nearly equal to n_3ek_3, so that Barnes and Bennett's measurements give k_3 and not k, because their method of calculation gives c_0/n_3e and not $c_0/(n_1 + n_2 + n_3)e$.

A. A. Noyes and the writer divided the conductivity by e and by the mobility given by the Hall Effect, which they took to be the mobility of the free electrons. If the Hall Effect really gives $\bar{k}$ and not k_3, their calculation of K would be invalidated unless k_3 and $\bar{k}$ are nearly equal.

It is easy to see that the Hall Effect should give the mobility of the free electrons k_3 when k_3 is much larger than k_1 and k_2. For the transverse field of the Hall Effect must be such as to prevent any appreciable transverse current or, since the current is practically carried by the electrons, it must prevent any transverse drift of the electrons relative to the other ions and the gas. This requires that $Hk_3X = Y$, where Y is the transverse field of the Hall Effect. Even if the electrons are only free part of the time, this makes no difference, because when they are not free they do not drift appreciably, so that the field Y must be large enough to prevent them drifting while they are free. The Hall Effect should therefore give the mobility of the free electrons.

The equation $Y = Hk_3X$ may also be obtained as follows: If i is the current density in the flame, then there is a force on unit volume equal to Hi. This force must be the force on the positive and negative ions due to the transverse field Y because there is no resultant transverse force on the electrons, so that $Hi = Ye(n_1 - n_2)$. But $i = en_3k_3X$ and $n_1 - n_2 = n_3$; hence $Hen_3k_3X = Yen_3$ or $Hk_3X = Y$. The theory of the Hall Effect in flames differs from that in solutions, where it is necessary to consider the concentration gradient set up by the action of the electric and magnetic fields, because in flames there is a rapid stream of gas across the space between the electrodes so that any particular metal atom or ion is only between the electrodes for a small fraction of a second. It seems, therefore, that in flames the condition which has to be satisfied is merely that the transverse current must be zero and not the more stringent condition that the transverse streams of each sort of ion present must each be zero.

There are strong reasons for believing that the fractions f and q in flames at about 2000° K. are constants independent of the concentration of the solution sprayed into the flame. The equations $b = Ac_0/Bfq$ and $a = A^2c_0^2/KBfq$, since b and a are found to be constants independent of the concentration, show that fq is constant.

It was shown by Gouy* in 1879 that the intensity of the light from a flame containing sodium is proportional to the square root of the product of the concentration of the solution of sodium salt sprayed into the flame and the thickness of the flame measured along the direction in which the light intensity is measured. The light intensity is therefore proportional to the square root of the amount of sodium per square centimetre in the flame and so for a fixed amount of sodium independent of the concentration in the flame. Now the sodium light is undoubtedly due to free sodium atoms, so that it follows from Gouy's results that the fraction of the salt reduced to the metal in the flame is independent of the concentration in the flame. To make this clear consider two flames one 5 cms. thick and the other 50 cms. thick, and suppose the concentration of the sodium, free and combined, in the 5-cm. flame 10 times that in the 50-cm. flame. According to Gouy's result both flames will give equal amounts of sodium light and must therefore contain equal amounts of free sodium atoms per square centimetre. Since they also contain equal total amounts of sodium, it follows that the fraction of the sodium in the free state is the same in both flames, although the concentration in one flame is 10 times that in the other. These results of Gouy's have recently been confirmed by a series of experiments done in this laboratory, an account of which will shortly be published.

---

* 'Ann. chim. et phys.,' vol. 18, p. 5 (1879).

We may conclude that the fraction f is a constant in sodium flames and so probably also in flames containing the other alkali metals. It may be supposed that any alkali metal salt is converted into hydroxide in the flame and that the hydroxide is reduced to the metal according to the reaction $2MOH + CO = 2M + CO_2 + H_2O$. This will make the fraction $f = M/MOH$ independent of the concentration, because the amounts of CO, CO_2 and H_2O present in the flame are very large compared with the amount of salt sprayed in.

If the fraction f is constant it follows that the fraction $q = n_3/(n_2 + n_3)$ must also be constant because, as we have seen, fq is constant. We should expect q to be independent of the concentration of the sodium or other alkali metal because the electrons are much more likely to get attached to atoms of the flame gases than to atoms of the metal or its compounds.

It is probable that q in flames at 2000° K. is nearly equal to unity, because Townsend* has shown that in gases at the ordinary temperature electrons remain free when their kinetic energy of agitation is about five times that corresponding to the ordinary temperature. In a flame at 2000° K. the energy of agitation of the electrons is nearly seven times that at the ordinary temperature, so we should expect the electrons to remain free.

The values of the mobility of the electrons 2450, 2600 and 2700 at about 2000° K. obtained by the writer,† J. S. Watt and Bennett respectively are smaller than might be expected for free electrons. The motion of electrons in gases at the ordinary temperature has been very fully investigated by J. S. Townsend and his co-workers, and a consideration of their results shows that a mobility of about 2600 for free electrons in a flame is not at all impossible. For example, it is found that the velocity of electrons in carbon dioxide is 120,000 cms./sec. when $Z/p = 0 \cdot 25$, Z being the electric intensity in volts per centimetre and p

---

* 'Motion of Electrons in Gases' (Clarendon Press, 1925).

† 'Phys. Rev.,' vol. 3, p. 375 (1914).

the pressure in millimetres of mercury. The kinetic energy of agitation of the electrons with $Z/p = 0 \cdot 25$ was $1 \cdot 2$ times that corresponding to the ordinary temperature. The density of the gases in a flame at 2000° K. is about that at 120 mm. pressure at the ordinary temperature. In CO_2 at 120 mm. the velocity due to 30 volts per centimetre would be 120,000, so that the mobility would be 4000 at the ordinary temperature. But in the flame the energy of agitation is at least $6 \cdot 7/1 \cdot 2$ times that at the ordinary temperature, so that the mobility would be only $4000/\sqrt{5 \cdot 5}$ to 1700 cms./sec. per volt/cm.

Townsend finds that in some gases the energy of agitation of free electrons may be considerably greater than that of molecules of the gas even in the absence of an electric field and that the velocity of the electrons due to small fields differs greatly in different gases. The gases in flames are mainly nitrogen, water vapour, carbon monoxide and carbon dioxide, and it is impossible to say what the mobility should be in such a mixture at 2000° K. It is, however, clear that a mobility of about 2600 is not impossible or even improbable. Townsend's results show that the mobility cannot be calculated even roughly by means of the various formulæ, based on kinetic theory, which have been proposed, because the mean free path and the energy of agitation vary greatly in different gases in a way which cannot be foretold.

Some measurements of the mobility of the electrons in flames have given considerably higher values than 2600. For example, E. Gold* got about 10,000, B. T. Barnes about 8000, and A. B. Bryan and the writer† got values from 4000 to 28,000. The method used by E. Gold is of very doubtful value, and the possible error in Barnes' determination is probably great enough to account for the difference between 8000 and 2700. A. B. Bryan and the writer calculated the mobility from the results of experiments on flames with high-frequency alternating currents. The theory of these experiments used in the calculation was admittedly only a rough approximation, and in any case there is no reason to suppose that the oscillations of the electrons, due to rapidly alternating electric fields, would correspond at all closely to the mobilities obtained with steady fields. In very rapidly alternating fields the energy of agitation may be less than in steady fields.

The relative conductivities of the alkali metals in flames at about 2000° K., as found by different observers, taking that of potassium equal to 10, are given in the following table. The last column contains the relative conductivities, according to Saha's theory, taking the conductivities to vary as the square roots of the equilibrium constants.

---

* 'Roy. Soc. Proc.,' A, vol. 79, p. 43 (1907).

† 'Phys. Rev.,' vol. 23, p. 195 (1924).

Arrhenius* (1891).		Smithells, Dawson,† H. A. Wilson (1899).	H. A. Wilson‡ (1915).	A. B. Bryan§ (1921).	B. T. Barnes‖ (1924).	Zachmann ¶ (1924).	Bennett** (1927).	Saha†† (1920).
Li	0·83	0·70	—	—	—	3·3	—	0·48
Na	1·65	2·08	1·1	1·2	—	7·5	—	1·00
K	10	10	10	10	—	10	—	10
Rb	16·3	19·7	12·8	—	—	—	(15·9)	15·9
Cs	24	38·6	27	17·5	(36·2)	—	33·1	36·2

* *Loc. cit.* † 'Phil. Trans.,' A, vol. 193, p. 89 (1899). ‡ *Loc. cit.* § *Loc. cit.*
‖ *Loc. cit.* ¶ 'Am. d. Phys.,' vol. 74, p. 1 (1924). ** *Loc. cit.* †† *Loc. cit.*

Barnes found the conductivity of cæsium to agree with Saha's theory, and Bennett found that of rubidium to agree also at 2000° K. This is indicated by the values in brackets. Bennett found the relative conductivities of cæsium and rubidium to be in the ratio 2·08, which makes that of cæsium 33·1, taking that of rubidium to be 15·9.

It appears that there is a rough agreement with Saha's theory, except in the case of Zachmann's values for sodium and lithium. It is well known that the conductivity imparted to a flame by potassium salts is much greater than that due to sodium salts, so that there seems to be no doubt that Zachmann's value of the relative conductivities of potassium and sodium is erroneous. A. B. Bryan's value for cæsium seems to be too small. The considerable differences between the relative values found by different observers show clearly the great difficulty of making such measurements accurately.

The following table gives the mean values omitting Zachmann's results and A. B. Bryan's result for cæsium.

——	——	Saha's theory.
Li	0·77	0·48
Na	1·5	1·00
K	10	10
Rb	16·2	15·9
Cs	31·8	36·2

The agreement seems to be sufficiently good to show that Saha's theory covers at any rate the main features of the process of ionisation in flames.

Summary of Conclusions.

(1) The assumptions made by A. A. Noyes and the writer—

(*a*) that all the metal is reduced to the free state,

(*b*) that all the negative carriers are free electrons,

in calculating the equilibrium constants of the reaction $M = M_+ + e$, where M denotes an atom of the alkali metal and e an electron, may be omitted without making any difference to the equation which they used.

(2) It is very probable that the fraction (f) of the metal reduced to the free state and the fraction (q) of the negative carriers which are electrons are both nearly equal to unity in flames at about 2000° K.

(3) In any case the fractions f and q are nearly independent of the concentration of the metal in the flame at 2000° K.

(4) The results obtained by different observers on the relative conductivities of the alkali metal vapours in flames at 2000° K. agree approximately with Saha's theory.

The Light of the Night Sky: its Intensity Variations when analysed by Colour Filter.—III.

By Lord RAYLEIGH, F.R.S., with the co-operation of others.

(Received February 23, 1928.)

§ 1. *Introduction. Methods used.*

This work has been steadily proceeding since the last publication upon it,* and the time has come to report the progress made up to the present.

It was felt to be urgent that the intensities and their variations should be compared in a series of observations at different stations scattered over the globe. This field of work was only slightly touched upon in previous papers, and the few provisional conclusions which were drawn have not been sustained by a more thorough investigation.

In this further work the spontaneous light of crystals of potassium uranyl sulphate has in all cases been used as a standard. Experience has led me to suspect that a mixture of this salt with radium as previously described is not to be relied on. On the other hand, the uranium salt by itself has given no reason for suspicion. Different specimens of the uranium salt differ among themselves, and, indeed, a bottle full of crystalline lumps of one sample photographed with a long exposure by its own light showed considerable patchiness. This is attributed in the main to varying transparency of the crystals, which, of course, alters the depth from which light can reach the surface. Experiments on keeping the crystals for months in an atmosphere saturated with water vapour, in a very hot room, did not show any discoverable change of intensity. Desiccation *in vacuo* with phosphorus pentoxide did seem to produce a diminution of light in the course of months; but this is very drastic treatment. In actual use the standards are sealed up, so that it is only in case of a failure of the cement that atmospheric influences could act.

---

* Previous papers :—'Roy. Soc. Proc.,' A, vol. 106, p. 117 (1924). Referred to as I. 'Roy. Soc. Proc.,' A, vol. 109, p. 428 (1925). Referred to as II.

There is no reason for suspecting loss of brightness in the periods which concern us here. The oldest sample, dating from 1908 or perhaps 1907, is quite as good as any of the several samples that I have. In fact, it is only equalled by a preparation made by slow crystallisation in 1925.

For the co-operative scheme the simple type of photometer figured in I, p. 122, was adopted, on account of the robustness and good travelling qualities.* These consist simply of a Lummer tube with the field divided into two vertical strips. The right-hand (transparent) is backed by the uranium salt. The left-hand part, silvered, reflects the sky, with a colour filter interposed. Choice from a series of neutral glasses allows the filtered sky light to be matched with the standard.

As in the earlier work three filters were used, one designed to isolate the green auroral line which is the most marked feature in the spectrum of the night sky, and the others designed to transmit the region (*a*) on the red side, and (*b*) on the blue side of the line, excluding the line itself in each case. These are called the red, auroral, and blue filters.†

The series of neutral glasses were calibrated with a rubidium photoelectric cell.‡ Advantage was taken of the lack of parallelism in the commercial sheets of neutral glass to select small pieces having a predetermined density. The density of each piece ($\log_{10}$ opacity) was chosen to have the value

$$0\cdot1, \quad 0\cdot2, \quad 0\cdot3 \ldots 0\cdot7,$$

with a tolerance in each case of $\pm\,0\cdot02$. The density was examined under each of the colour filters with which it would be eventually used. The value was not the same with each coloured light, owing to deficient neutrality of the "neutral" glass. But it was found possible to provide glasses having densities within the tolerance limits stated for all the colours concerned. The book-keeping involved in recording observations was thus greatly economised.

---

* On other grounds the more complicated design illustrated on p. 124 is still preferred.

† They are described in detail in I, pp. 119-120.

‡ II, p. 429.

It would have been an advantage if all the instruments could have been exactly comparable, but it was not found feasible to make them so, on account of the variable quality of the uranium source, and also to some extent of the colour filters, which did not all come from the same batch of material.

It became necessary, therefore, to intercompare the instruments before issuing them. Previous experience had shown that it was very difficult to do this satisfactorily by comparisons made on the actual sky, and a different method was adopted. "An artificial sky" was arranged by feebly illuminating an opal glass by a small incandescent lamp, controlled by resistances at the observer's hand.

The photometer under test was placed with its sky aperture up to the opal glass, without any colour screen, and the luminous intensities equalised. When the adjustment was satisfactory the observer called out to an assistant in the next room, who read the voltage across the lamp terminals on a conveniently illuminated voltmeter, and recorded the result. It is a great advantage of this method that the adjustment can be made in a dark room, and readings taken and recorded in a lighted one. Thus, a great number of readings can be made in a moderate time, and bias is largely eliminated.

In this way the voltage to match the source in each instrument was determined.

My own original instrument labelled 0 was adopted as a master standard. It was brighter than any of the later ones, and accordingly gave a higher voltage reading. The light from the master standard was cut down in a known ratio with a neutral glass, previously calibrated for opacity by the photoelectric cell, and the lowered voltage on the "artificial sky" again determined. It will readily be understood that the brightness corresponding to any intermediate voltage could be interpolated, and the various instruments compared with the standard in this way.

It remained, however, to take the colour filters into account. These could not be trusted to be all absolutely identical, as was found by tests with a photo-electric cell. The transmission of each was therefore tested in comparison with the filter of the same colour belonging to the master standard instrument. In this way a correction was found for each filter, and combined with the correction appropriate to the instrument it belonged to, to form a total correction for red, auroral, blue, respectively.*

The scale now used for recording and discussing the results is no longer an intensity scale as in the former work, but a magnitude scale, *i.e.*, one in which each step is a constant multiple of its predecessor. This is what we get observationally in the first instance, and it appeared on consideration that nothing was gained by translating the results into intensities before discussing and comparing them, as was done in the earlier work (I and II). The geometric scale is in many respects more convenient.

The zero reading is that which gives a match without one of the neutral glasses. A reading of (*e.g.*) + 3 means that the third neutral glass has to be used *over the sky.* A reading is recorded as − 3 if the same glass has to be used *over the uranium source.* Intermediate interpolated values are entered as, *e.g.*, 3·5. The zeros for the three separate regions of the spectrum, red, auroral, blue, are related in an arbitrary manner, though the factor involved in passing from one neutral glass to the next is, of course, the same for each, so far as the glasses deserve the name neutral. Within the practical limits of experimental error, they do deserve it.

This is all that is done by each local observer in obtaining and recording his results. They are reduced to the standard scale at headquarters by applying a subtractive correction, which gives the result which would be obtained under the same conditions on the master standard instrument, with the filters belonging to the latter. The scale numbers thus adopted are those of the neutral glasses, thus the intensity is multiplied by passing up one unit by the antilogarithm of 0·1 or 1·259. Three steps on the scale are equivalent to a factor of $(1 \cdot 259)^3$ or by approximately a doubled intensity. This is a convenient rule to remember. Ten units on this scale is equivalent to one unit on the scale of densities used by Hurter and Driffield in their photographic researches. It was convenient to adopt this modification of their scale, so as to make the actual set of neutral glasses proceed by steps of one unit, fractions being used only for interpolation.

---

* The reader may wonder why this apparently circuitous procedure was adopted instead of testing each instrument on the artificial sky with, *e.g.*, the red filter in position. The answer is that the independent comparison of the filters with the photoelectric cell is quick and accurate, while the visual comparison of the instruments with the artificial sky is much more exacting and requires many repetitions. Hence the objection to multiplying threefold the number of the latter class of measurements.

§ 2. *Stations and Observers.*

For the observations in various parts of the world I have been able to rely on the kindness of scientific friends who either undertook the work themselves or were able to find other capable observers known to them who were so kind as to undertake it. In the choice of stations, I was anxious as far as possible to avoid the neighbourhood of large towns, which introduce the disturbing element of artificial light. Astronomical observatories when so situated have an obvious advantage from the presence of a staff of skilled observers who are normally on night duty. I was able to negotiate in many cases at the Astronomical Union at Cambridge in the summer of 1925, and to hand over the instruments. These were constructed in my own laboratory by Mr. R. Thompson.

I now give a list of stations from which observations were taken in order of longitude, west to east, with brief notes as to the conditions of observation and the observers.

Hawaii (Lat. 19° N., Long. 155° W.).—Mr. Babcock, of the Mount Wilson Observatory, arranged with Mr. H. E. McComb, of the U.S. Magnetic Observatory, Ewa, Oahu, Hawaii, to make observations, which the latter kindly undertook. The nearest village is Ewa, two miles distant, and no others within 10 miles. Mr. McComb considers that the Ewa lights had no effect.

Victoria, British Columbia (Lat. 40° N., Long. 123° W.).—The observations were made at the Dominion Astrophysical Observatory by Dr. Plaskett and his assistants, Mr. Harper, Mr. H. H. Plaskett, and Mr. Pearce, which ever happened to be on duty. Dr. Plaskett found good agreement between the various observers when intercomparisons were made. I understand that the north sky is reasonably free from interference from town lights.

Auroral displays are frequent at this station. This aspect of the subject will be discussed in a later part of the paper.

Mount Wilson, California (117° W.).—The instrument was taken over by Mr. Babcock, who, it may be remembered, kindly assisted me earlier with observations from his own house in Pasadena, California. On comparing the observations made at sea with those made afterwards at his own house on the outskirts of Pasadena, he found that there was serious interference from town lights. It is necessary, therefore, to reject the conclusions drawn from the former observations made at Pasadena (II, pp. 441–42).

Mr. Babcock arranged for observations to be made by Mr. Humason on Mount Wilson, where, of course, the atmospheric and other conditions are ideal.

Pomona College, Claremont, California (117° W., 30 miles east of Pasadena).—Mr. Babcock kindly arranged with Prof. Brackett to make observations at this station, which is fairly free from disturbance of town lights. The results show about the same range as other stations, but the values are about 3 units higher than, *e.g.*, at Mount Wilson. This anomaly will be discussed later.

Pinehurst, N. Carolina (Lat. 36° N., Long. 79° W.).—Prof. J. C. MacLennan, F.R.S., kindly made a few observations for me from this station. They refer to two days only, and, judging from the map, should be free from disturbance of artificial light.

Kingston, Ontario (Lat. 44° N., Long. 77° W.).—A few readings were received from this station through the kindness of Prof. J. C. McLennan, F.R.S. He states that they were taken far enough from city lights to avoid any interference from these. Display of aurora is frequent at this station.

Arequipa, Peru (Lat. 16° S., Long. 71° W.).—Through the kindness of Dr. Harlow Shapley, director of the Harvard Observatory, observations were made at this station by Dr. J. S. Paraskevopoulos. The site is, I believe, quite free from disturbance of artificial lights. The altitude is 8050 feet.

Lerwick, Shetland Islands (Lat. 60° N., Long. 1° W.).—By the kindness of Dr. G. C. Simpson, F.R.S., director of the Meteorological Office, Air Ministry, observations were made at this station by Mr. A. W. Lee. This is the farthest north of the stations, and of course only admits of observations in the winter. The same problem of auroral displays arises as in the case of Victoria, B.C., and was dealt with in the same way. At this station observations were usually taken both north and south, but there was rarely any difference except in the case of a brilliant display. Polar aurora is usually limited to a lower elevation in the north than that observed.

At the Lerwick Observatory there is no interference from artificial lights. The town is some miles away, and comparatively small and dimly lighted with gas.

England.—My own observations have been continued at Terling, Essex (Lat. 52° N., Long. 1° E.) as before. During a part of August and September in each year I have resided in Northumberland, near Hexham (Lat. 55° N., Long. 2° W.), and the work had to be carried on there instead of at Terling. In view of the general results of the investigation extending over enormously wider ranges of latitude, I have not thought it necessary to distinguish between these stations, and they are classed together as "England." Both positions are perfectly free from suspicion of interference by town lights.

Observations at this station extend from the year 1923, and are in this respect of special value for testing questions of periodicity. Unfortunately some of the earlier observations of the red and blue components are now suspect, because the standard of light was the mixture of uranium salt with radium, which is believed not to be of trustworthy constancy. All of those before August, 1924, have been rejected. From that date I have an independent check on them. The observations with the auroral filter have fortunately been made with the same instrument from the very beginning—this is the instrument now used as a master standard (No. 0). They form, therefore, an unbroken series of comparisons of the sky with a single uranium standard, made throughout (with rare exceptions) by one individual, the present writer.

The scale now used is the geometrical one above explained, and the older results are *restated on that basis*. This must be borne in mind in referring to the former papers I and II.

Cape of Good Hope (Lat. 34° S., Long. 18° E.).—H.M. Astronomer, Dr. H. Spencer Jones, kindly promised to make observations, and has been able to provide a very complete series, but little interrupted by rainfall. Most of these he has made personally, but from June, 1927, onwards a series have also been made by Mr. R. Woodgate, of the observatory staff. Two instruments have been in use since that date. The situation of the observatory is not far enough from Cape Town to be altogether ideal for the purpose, and Dr. Spencer Jones thinks that they may have been affected to some small extent by the town lights. The large number and continuity of the observations, however, give special value to this series. Dr. Spencer Jones also made some observations while travelling elsewhere in South Africa.

Gilgil, Kenya Colony, E. Africa (Lat. 0°, Long. 37° E., roughly).—A few observations were kindly made for me at this station by my relative the Hon. Mrs. G. Cole. The site is far from town lights. Unfortunately owing to illness and other unfavourable circumstances the observations are limited to the month of September, 1925, after which Mrs. Cole was not able to continue.

Kodaikanal Observatory (Lat. 10° N., Long. 77° E.).—Height above sea-level 7600 feet. The director, Dr. T. Royds, kindly undertook to have the work done, and superintended it himself. The actual observations were taken by A. A. Narayana Ayer and P. R. Chidambara Ayer, Dr. Royds and S. S. Ramaswamy Ayyangar.

Canberra, Australia (Lat. 35° S., Long. 149° E.).—Dr. W. G. Duffield, the director of the Commonwealth Solar Observatory, kindly undertook to have the observations made. They have chiefly been carried out with his supervision by Mr. A. L. Kennedy, though Dr. Duffield himself has occasionally taken part.

This series, like those of the Cape, have the great merit of continuity, with little interruption from bad weather. The earlier part of the series, up to the middle of October, 1926, were taken at Hotel Canberra in the plain, where, however, I understand there are too few artificial lights to have a perceptible effect on the sky. Since the date mentioned they have been made at the observatory on Mount Stromlo, which is doubtless an ideal situation.

Christchurch, New Zealand (Lat. 44° S., Long. 173° E.).—The instrument was kindly taken out by Sir Ernest Rutherford, P.R.S., in the summer of 1925, and he left it in the hands of Prof. Farr, who kindly arranged with Mr. P. W. Glover to use it. Most of the observations were made about 2½ miles away from the town lights. This is nearer than might be desired and possibly some slight interference may be suspected.

Voyage from Marseilles to Boston.—Mr. Babcock, of Mount Wilson Observatory, made a series of observations on this voyage in August, 1925. The range of positions was from Lat. 39°, Long. 0°, to Lat. 41°, Long. 67° W.

Voyage to Melbourne.—The instrument destined for Canberra was taken out by Miss Natalie Allen, of the Physics Department, Melbourne University, who very kindly made observations from the bridge and from a cabin porthole, which she believes were not seriously interfered with by ship's lights. The observations range from Lat. 12° N., Long. 49° E., to Lat. 35° S., Long. 117° E. They were made in September, 1925.

§ 3. *Comparison of the Intensities observed and of their Extreme Ranges at Different Stations.*

In comparing the observed intensities at different stations, we are met by the difficulty that the number of observations available, and hence the representative character of the mean, varies very much at the different stations. Moreover, owing to unavoidable difficulties of organisation, the periods over which they extend are not by any means identical.

In the following table, a mean is given for each station and with it the months over which the observations used in deriving it extend, with the total number of observing nights therein. For the majority of stations, these means are derived from all the observations available. But for three stations—England, the Cape, and Canberra (Australia)—the observations have gone on much longer than elsewhere, and I have arbitrarily selected a period of about a year, which roughly coincides with the period of observation at most of the other stations.

The particular stations named, even when so treated, afford an ample number of nights for a representative mean value, and some approximation to simultaneity with the other stations seems an advantage.

Station.	Range in time.	Total number of observation nights.	Mean values.		
			Red.	Auroral.	Blue.
Hawaii	Oct., 1925, to Nov., 1926	116	−3·2	+0·3	+6·5
Victoria	Sept., 1925, to Mar., 1927	108	−3·2	+0·8	+6·3
Mount Wilson	Sept., 1925, to Feb., 1927	83	−2·9	+1·1	+6·8
Claremont	Oct., 1925, to Feb., 1926	57	−0·7	+2·1	+8·9
Pinehurst	January, 1926	3	−2·1	+1·7	+5·9
Kingston	April, 1926, to May, 1926	6	+0·1	+2·3	+9·5
Arequipa	April, 1926, to Nov., 1926	31	−3·1	−0·3	+6·5
Lerwick	Sept., 1925, to April, 1927	86	−2·8	+0·6	+5·7
England	Nov., 1925, to Oct., 1926	81	−2·5	+0·5	+6·4
Cape	Nov., 1925, to Nov., 1926	199	−2·7	+0·7	+7·1
Gilgil	September, 1925	5	−2·7	+1·5	+6·6
Kodaikanal	Oct., 1925, to Jan., 1927	54	−2·8	−0·1	+6·2
Canberra	Nov., 1925, to Oct., 1926	149	−3·2	+0·4	+6·5
Christchurch	Feb., 1926, to Feb., 1927	56	−1·0	+0·4	+7·0
Voyage to Boston	August, 1925	6	−3·2	+0·8	+6·0
Voyage to Melbourne	September, 1925	10	−3·3	+0·0	+5·6

Coming now to the question of the amount of variation, we have the same difficulty that the material is so much less abundant at some stations than at others. When the material is ample, and when the measurements are of all desirable accuracy, the standard deviation is of course preferable to the extreme range as a measure of dispersal about the mean. In measurements, *e.g.*, of the barometric height, the observation error is in general very small compared with the deviation of an observation from the mean. But in the present investigation these conditions unfortunately do not prevail. It is only in the more extreme cases that the deviation from the mean is a large multiple of the probable error of observation. Although extreme cases weigh heavily in determining the standard deviation, yet it is judged that in the present case they do not weigh heavily enough, and the extreme range is considered more instructive. Moreover, in the cases where the observations are too few to be reasonably representative, the extreme range rather than the standard deviation makes the most of the meagre information we have.

Fig. 1 shows graphically the mean values of the previous table, which are marked by crosses, with the extreme ranges in each component* marked by the

---

* In marking the extreme ranges I have used all observations to date, without limiting the values for England, the Cape and Canberra to the particular range of months used in deriving the means, see above, p. 18. This has led to a slight extension of the range in certain cases.

Throughout this section the observations made during auroral displays have been excluded. See § 6.

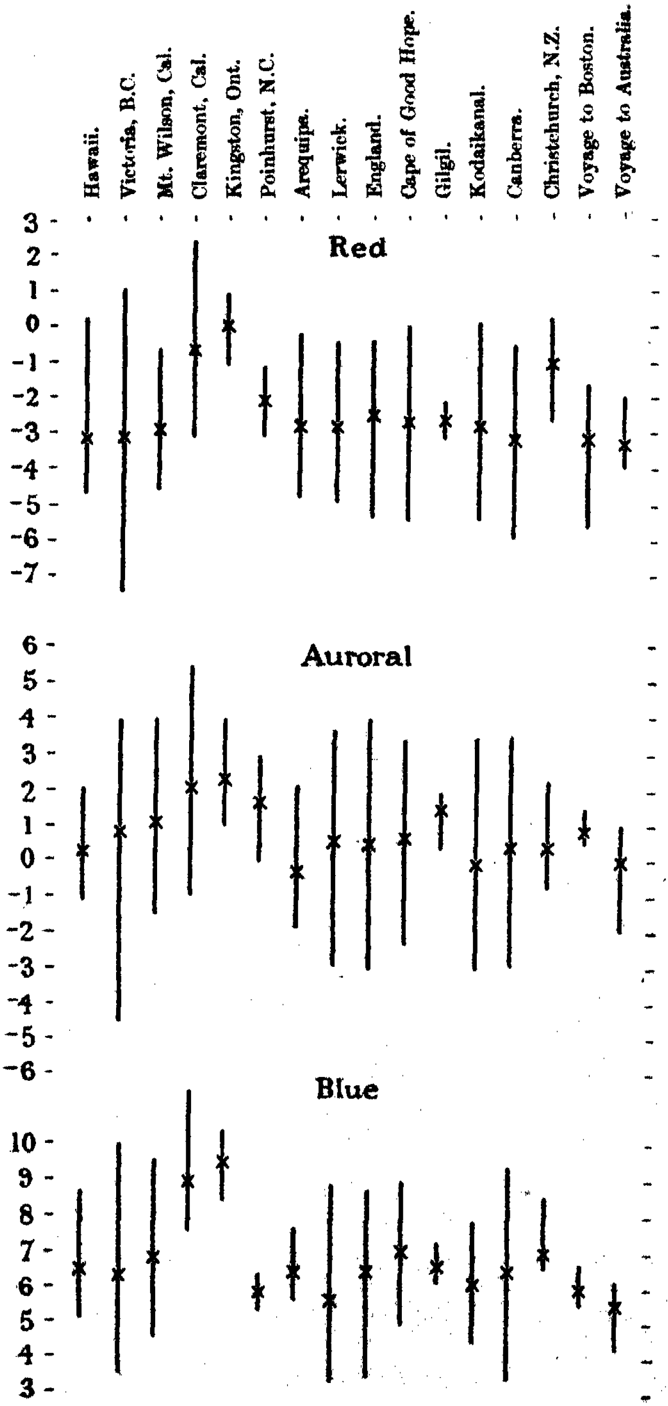

Hawaii.
Victoria, B.C.
Mt. Wilson, Cal.
Claremont, Cal.
Kingston, Ont.
Poinhurst, N.C.
Arequipa.
Lerwick.
England.
Cape of Good Hope.
Gilgil.
Kodaikanal.
Canberra.
Christchurch, N.Z.
Voyage to Boston.
Voyage to Australia.
Red
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
Auroral
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
Blue
10
9
8
7
6
5
4
3

FIG. 1.

vertical lines. As before emphasised, the zero for blue is an arbitrary value having an arbitrary but constant relation to the zero for either of the other components.

The general conclusions to be drawn from this diagram appear to be as follows:—*Firstly*, fairly normal values can be stated for the intensity of each component at any part of the world. These values are somewhat as follows, on the various arbitrary scales:—

Red..................	—2·5
Auroral	+0·8
Blue	+6·5

The only case when these values lie appreciably outside the observed range is Kingston, Ontario, but this may be explained by the extreme paucity of the observations which (it may reasonably be supposed) happen to be somewhat abnormal.

There are, however, two other stations, Claremont, California, and Christchurch, New Zealand, where anomalies occur.

At Claremont the observed readings are about 3 units higher (double) than those at other places, though the range is about the same. It is hard to avoid the suspicion that some cause of error has crept in. The station is thought to be free from disturbance by town lights, and I am inclined to attribute the anomaly to a derangement of the apparatus. The cell which contains the uranium salt had certainly got turned round by the jarring of travel when it came back into my hands, no longer presenting the proper face as when calibrated. The back of the uranium salt was packed into the cell with glass wool, which would diminish the light and therefore increase the apparent brightness of the sky. The tubular glass cell was apparently pretty tightly packed in its place, both originally and when returned, but, nevertheless, rotation had occurred in the meantime. If the observations were made in these conditions the anomaly would be explained.

If we regard the exceptional brightness at Claremont as real, we should be compelled to attribute it to some exceptional local cause. But at Mount Wilson, only 30 miles distant, the brightness is normal.

The other marked anomaly is Christchurch, New Zealand, where the mean value of the red component is about 1·5 units higher than at other places. In this case, too, I should be inclined to seek some observational cause, but it can hardly be of the same character as in the previous case, since only the red readings are exceptionally high. It is possible that town lights produced more effect than was supposed. It is perhaps also possible that some physiological peculiarity of the observer's vision enters. It has been assumed that the comparative luminous intensity as a function of the position in the spectrum is the same for all observers, when vision is dark-adapted, and it is on this basis that the filtered sky light is compared with the light of the uranium salt, which has in general a remarkably different spectral range. So far as my own experience with occasional assistants goes, this assumption is justified, and Dr. Plaskett, too, has found that his observations agree with those of his assistants at the Victoria Observatory. It is, however, possible that cases exist which depart from the normal. Any marked selective absorption in the tissues of the eye would certainly produce such an effect. The question deserves further investigation.

Secondly.—At all stations where a reasonably large number of nights have been observed, the intensity in each component varies over a considerable range. These ranges are of the same order of magnitude in each of the components, but they tend on the whole to be largest for the auroral component, and smallest for the blue, the red being intermediate.

As a typical example the observations of Dr. Spencer Jones at the Cape may be cited. They give for the range

$$\left.\begin{array}{ll}\text{Red} \ldots\ldots & 5\cdot5\\ \text{Auroral} \ldots & 5\cdot7\\ \text{Blue} \ldots\ldots & 4\cdot0\end{array}\right\}\text{ Equivalent to a factor of }\left\{\begin{array}{l}3\cdot5\\ 3\cdot7\\ 2\cdot5\end{array}\right.$$

The largest range observed in any component is for the auroral component at Victoria, which is 8·5 scale units, representing a factor of 7·1. This observatory is within the zone of frequent polar auroras, and therefore the above estimate involves the criterion for exclusion of the latter. This is pretty definite, as explained in § 6, and in fact the large range at Victoria is not due to values exceptionally *high* at the top of the range, but to values exceptionally *low* at the bottom.

§ 4. *Are the Variations at Different Stations correlated?*

The variations of intensity which form the subject of this investigation do not occur uniformly all over the world. They are conditioned in large part at any rate by local circumstances. To illustrate this, some striking illustrative cases will be given before discussing the subject by statistical methods.

Date.	Place.	Red.	Auroral.	Blue.
January 16, 1926	England	−4·4	−1·4	+5·8
	Cape	−1·9	+2·0	+9·0
March 2, 1926	England	−3·6	−0·2	+6·4
	Cape	−0·4	+2·0	+7·7
April 15, 1926	England	−3·6	−0·8	+6·4
	Cape	−2·4	+2·4	+8·3
September 19, 1925	England (Northumberland)	−3·6	+1·7	+6·4
	Shetland (Lerwick)	−4·4	−2·8	+3·4
June 7, 1926	Hawaii	−4·7	−0·9	+5·1
	Canberra	−2·4	+2·1	+7·9

These cases have been chosen to show contrast. The *mean* values at the two places are in each case very nearly the same. Yet we see that occasion may be found where the intensity at one is double or more than double that at the other.

The earlier investigation* showed that there was a strong correlation between the red and auroral intensities on a given occasion, and a rather less strong but still marked correlation between these and the blue. The present data generally confirm this conclusion, and I have not thought it necessary to elaborate it further by strict statistical methods. What follows therefore will for the most part be confined to the auroral values. The reason for selecting these is that they are more accurate than the red on account of the greater intensity of light, and that the observed variations are over a larger range than either of the other components. Moreover, the intensity of a definite line in the spectrum, as approximately isolated by the auroral filter, has obvious attractions as a subject of study, which the less well-defined blue and red regions (mainly of continuous or band spectrum) do not possess.

It will be shown later that there is pretty strong evidence of an annual periodicity at any rate in England, which is in much higher latitude than the generality of stations. Nearer the equator this periodicity would become less important, and within the tropics, if there were any perceptible effect dependent on the sun's motion in declination, it would no doubt be semi-annual. In fact, however, the evidence from available data of an annual effect is not strong except for England, where the effect itself is probably much larger and the length of time over which the observations extend is much greater.

In latitudes high enough for such a periodicity to be apparent a positive correlation would be implied between stations in the same hemisphere, and a negative one between stations in the opposite hemisphere. The annual periodicity, however (supposing the evidence for it to be admitted as sufficient), certainly does not account for all the variations. There is certainly an irregular variation superposed on the regular annual periods, as, *e.g.*, in the case of temperature records, and the question arises whether this is correlated at stations widely distant. If the variations were due to cosmic causes such, *e.g.*, as outbursts on the sun, a positive correlation would be expected.

---

* II, p. 435.

Unfortunately the material available for such a discussion is very limited. It is necessary to work with stations not separated in time by more than a few hours, and to have not less than about 50 nights on which observations are available at both places. The stations were in fact chosen widely separated in longitude, without enough regard to the first consideration, and the vagaries of weather at two separate stations make the latter condition particularly hard. The only cases where both conditions are reasonably fulfilled are :—

(*a*) Terling and Lerwick, Shetland.
(*b*) Terling and the Cape.
(*c*) Canberra and Hawaii.

In the case of Terling and Lerwick, 45 nights of simultaneous observation were available, September, 1925, to April, 1926, inclusive. These gave :—

	Terling.	Lerwick.
Mean intensity	+0·5	+0·6
Standard deviation	1·1	1·1
Correlation coefficient..........	0·09±0·10	

Here there is clearly no evidence of significant correlation.

In the case of Terling and the Cape, discussion of the auroral intensity on 53 nights, October, 1925, to April, 1926, inclusive, give for the

	Terling.	Cape.
Mean intensity	+0·4	+0·5
Standard deviation	1·0	0·93
Correlation coefficient	−0·24±0·09	

It will be seen that the correlation is negative though too small to be of much significance. A negative coefficient would result from the annual effect taken alone. There is no evidence here of a positive correlation of the irregular variations.

For the case of Canberra and Hawaii, all the auroral intensities are available at both for 87 nights, November, 1925, to October, 1926. These have all been used.

	Canberra.	Hawaii.
Mean intensity	+0·2	+0·3
Standard deviation	1·2	0·67
Correlation coefficient	+0·15±0·07	

Here the correlation, such as it is, is positive, but again too small to be of definite significance.

§ 5. *Periodicities.*

We have seen in detail that the light of the night sky is of variable intensity, the variations being largely irregular. As a step towards gaining some knowledge of causes, it is desirable to enquire whether anything of the nature of a periodicity or of a secular variation can be detected when the irregular variations are averaged out. If secular variation is disclosed, it will doubtless represent a periodicity with periodic time considerably longer than the time over which the observations extend.

With the data available discussion may be limited to the question of :—

(1) Annual periodicity.
(2) Variation in the period of the solar cycle.

An annual periodicity was early suspected in my own observations in England.* At the time of writing, observations lasting over five years are available, so far as the auroral component is concerned. To examine impartially whether or not an annual variation is present, the observations of each calendar month (November, 1926, *e.g.*) are averaged, and the mean is adopted as representative of that particular month, without further reference to the data for individual nights.

These mean values, or the general mean for the corresponding months (*e.g.*, November, 1923, 1924, 1925, 1926, 1927) are plotted against the ordinal number of the month (1–12). Owing to the lack of data for the summer months, May, June, July, when residual daylight prevents the observation, the data are arranged from August to April inclusive.

---

* II, p. 442.

	August.	September.	October.	November.	December.	January.	February.	March.	April.
1922–3....	—	—	—	—	—	—	−0·5	0·0	−1·9
1923–4....	+0·0	+0·3	+0·7	+0·0	−2·2	−2·2	−1·6	−0·3	−1·7
1924–5....	+1·0	+0·7	+1·6	−0·9	−0·8	−1·5	−1·2	−0·8	−1·8
1925–6....	+0·8	+0·9	+0·8	+1·5	+0·0	+0·5	−0·3	−0·4	+0·0
1926–7....	+0·5	+0·8	+1·3	+1·1	−0·4	−0·7	+0·0	+0·8	−0·5
1927–8....	+2·0	+1·2	+1·8	+0·9	+1·0	−0·3	—	—	—
Sum	+4·3	+3·9	+6·2	+2·6	−2·4	−4·2	−3·6	−0·7	−5·9
Mean	+0·9	+0·8	+1·2	+0·5	−0·5	−0·8	−0·7	−0·1	−1·2

The individual seasons and the mean are plotted in fig. 2.

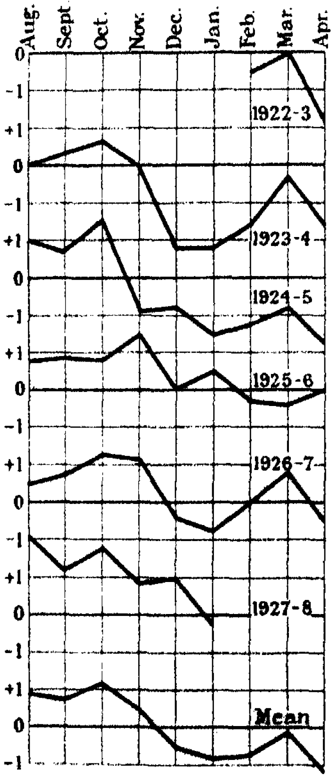


Fig. 2.

Although the results are not numerous enough to give complete statistical regularity, yet I think it is difficult to resist the indications of a definite annual period. The great fall of intensity usually occurs between November and December. In each of the four complete observing seasons (August to April) the mean of all months before this date is much greater than the mean after it. This fall must have its counterpart in a recovery, which would naturally be taken to occur six months later, *i.e.*, between May and June.

No observations can be made at that season. But in every one of the five cases available a considerable recovery has occurred at some time between April and August, in the interval when residual daylight prevents observations.

Special attention is directed to the mean curve at the bottom of fig. 2, which sums up the evidence for annual periodicity. The amplitude of annual variation indicated is about 2 units, or a factor of 1·6. The subordinate maximum in March is probably not significant, but due to coincidence of irregular variations.

Turning now to the second question, that of variation in the cycle of the sunspots, it is to be noted that the observations begin at about the time of sunspot minimum in 1923, and that according to present expectation the maximum will be reached in the present year, 1928. By analogy of other phenomena known to depend on the spot-cycle, we should anticipate an increase in the light of the night sky during this period, if such an influence exists at all.

To eliminate any effect of the annual periodicity just discussed, it is necessary to compare the intensity in successive years at the same season. If we look at any given column in the table (October, *e.g.*) we find a tendency to increase with time, though this tendency is not strong enough always to assert itself against the irregularities after the lapse of a single year. But, in every one of the nine columns, the last entry is substantially higher than the first. If the variations when cleared of the annual periodicity were quite unsystematic, the chance of this occurring would be only 2^{-9} or 1 in 512.

For greater statistical regularity, we may take the mean of all the months in a given observing season, or in a given year. To include observations taken in the spring of 1923, and in the present uncompleted observing season, the latter course is adopted. We then get

Year.	Sum.	Mean.
1923	$-3 \cdot 6$	$-0 \cdot 45$
1924	$-4 \cdot 2$	$-0 \cdot 47$
1925	$-1 \cdot 3$	$-0 \cdot 14$
1926	$+3 \cdot 1$	$+0 \cdot 34$
1927	$+6 \cdot 5$	$+0 \cdot 72$

Results for January, 1923, are lacking, as the work was only begun in February. According to the observed annual variation (fig. 2) this would tend to make the mean for 1923 too high, but in spite of this 1923 is nearly the lowest year. Each of the succeeding years shows an increase on the mean of its immediate predecessor. The result is shown graphically in fig. 3.

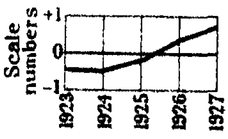


FIG. 3.

The mean annual increase is $0 \cdot 3$ unit, representing a factor of $1 \cdot 07$. Thus the results clearly indicate a general yearly increase of the intensity during practically the whole period of observation, relative to the uranium standard. This could be explained away if we assumed a slow loss of brightness in the standard, though it would be impossible thus to explain the annual variations, which show a diminution of sky intensity alternating with the increase.

Though there is no rigorous independent check on the uranium standard, yet I do not believe that this is the explanation. The uranium standard used throughout is the old specimen which has now been in my possession 20 years, and appears to be about as bright as at first* so far as my recollection goes. The general reasons that led me to believe in its constancy have already been given (p. 12).

The question remains whether the observed increase of sky brightness is connected with the solar cycle or not. All that one can say at present is that it is in general harmony with it over nearly half a period, from minimum to maximum; but at least a whole period is required for a decisive result. I hope to be able to continue for the additional six years, or even longer. If the expected decline of average brightness occurs, it will settle any possible doubts about a change in the uranium standard as an effective cause of the increase observed up to the date of writing.

Returning now to the question of annual periodicity, it seems probable that this depends on the sun's motion in declination. If so the effect should be negligible at the equator, and should show itself in the opposite phase in the southern hemisphere. The question arises whether the available data give any evidence of this.

The very full series of observations from the Cape (lat. 34° S.) and from Canberra (35° S.) give the best material. The season 1925–26 shows the expected effect to a marked extent, but the season 1926–27 cannot be said to do so. I have used the combined data from both places in averaging for each month. See table and fig. 4, which shows a marked maximum in April, in agreement with the expected phase.

Values through Auroral Filter, Southern Hemisphere.

	August.	September.	October.	November.	December.	January.	February.	March.	April.	May.	June.	July.
Cape, 1925–6	—	—	+0·0	+0·6	+0·4	+1·2	+0·5	+0·6	+1·4	+1·6	+0·1	+0·2
Cape, 1926–7	−0·2	+0·6	+0·5	+1·2	+0·3	+1·0	+0·3	+1·0	+0·2	+0·4	−0·8	−0·5
Cape, 1927–8	+0·0	+0·4	+1·3	—	—	—	—	—	—	—	—	—
Canberra, 1925–6	—	—	—	+0·3	−0·1	+0·8	+0·2	+0·9	+3·0	+1·6	+1·0	+0·3
Canberra, 1926–7	−0·8	−0·6	−0·8	−0·1	−1·5	−0·6	−0·8	+0·0	+0·2	−0·5	−1·2	−1·4
Canberra, 1927–8	−1·0	−0·8	—	—	—	—	—	—	—	—	—	—
Sum	−2·0	−0·4	+1·0	+2·0	−0·9	+2·4	+0·2	+2·5	+4·8	+3·1	−0·9	−1·4
Mean	−0·5	−0·1	+0·2	+0·4	−0·1	+0·6	+0·1	+0·6	+1·4	+0·8	−0·2	−0·3

---

* If we assume an exponential law of decay of the brightness of the uranium standard adjusted to explain the observed apparent increase of sky brightness, then 20 years ago the uranium salt would have been about four times brighter than it is now, and four times brighter than any specimen I can now obtain. I feel sure it was not so in fact.

At the same time I do not feel any great confidence that further continued observations will confirm the existence at these stations of so marked an annual effect as the diagram suggests.

The stations it is to be noticed are not nearly so far to the south as England is to the north of the equator. Indeed, the situation of England (52° N.) appears to be an ideal compromise. If we go farther still to the north, the part of the year during which observations can be taken is unduly restricted.

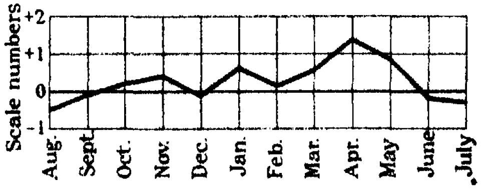


FIG. 4.

In this discussion of periodicities I have worked throughout with auroral intensities rather than red or blue. The latter give generally similar results.

§ 6. *Polar Aurora.*

In I, pp. 135–36, the distinction was emphasised between the *polar aurora* occurring typically in the auroral zones, and the *non-polar aurora* occurring all over the earth.

Both have the green line λ 5578, which has been shown by the work of McLennan and his colleagues to be due to oxygen, as the strongest feature of the visual spectrum. It may seem an arbitrary procedure to distinguish two separate phenomena when there is so much in common, and I know from correspondence and otherwise that some scientific men are unwilling to admit it. Nevertheless, after further experience and study, I am more than ever convinced that it is necessary.

The grounds on which this distinction was originally made were that the nitrogen bands (negative bands) occur only in the polar aurora ; that the polar aurora has a distinctive distribution in latitude which the other has not ; that the polar aurora often shows distinctive forms such as arcs and draperies, whereas the non-polar aurora is uniform ; while finally the polar aurora varies enormously in intensity in the course of a few minutes, whereas the non-polar aurora is often sensibly constant for days. I wish now to discuss the subject from the standpoint of intensity measurements.

We find, in the majority of stations, where " auroral displays " rarely if ever occur, that the range of intensity observed through the auroral filter, which I take as a measure of the intensity of the green line, has a fairly definite mean value of about 0·4 on my scale, with a standard deviation of 1·0. It is true that this intensity is subject to a yearly, and probably also to an 11-yearly, variation. But these variations are comparatively slight, and for the present purpose may be ignored.

Lerwick (Shetland), Victoria (British Columbia) and Kingston (Ontario) are stations where auroral displays are comparatively common.

At the Cape, or at Canberra, or Hawaii, they are practically unknown.

The case of England is intermediate. Auroral displays occur occasionally though not often.

Let us consider what happens in this intermediate case, where the statistical issue can be most easily presented. In the year 1926 my observations in England on 75 nights, in the absence of an auroral display, give for the auroral intensity :—

Range	−2·2 to +3·0
Mean value	+0·2
Standard deviation	1·1

Now only one auroral display was observed when it was found that

Auroral display gave	+16·0
Deviation from the mean	+15·8

thus giving an isolated deviation of no less than 14·4 times the standard. As is well known in all ordinary cases of statistical distribution, values three times the standard are nearly the highest that occur. A value 14·4 times the standard is fantastic, and shows that we are attempting to classify in the same scheme values that are not properly comparable at all. For a Gaussian distribution values even six times the normal only occur once in 10^9 cases. Values such as 14·4 times the normal are too large to come within the range of existing tables of the probability integral ; but to show what such values imply, an illustration from a more familiar field may be helpful.

For the height of the barometer at sea-level we may take as rough illustrative values :—

Range	28·5 to 31·0 inches
Mean	30·0 inches
Standard deviation	0·33 inch

Thus a positive deviation of 14·4 times the standard would imply

Barometric height........$30+14{\cdot}4\times0{\cdot}33 = 34{\cdot}7$ inches.

What conclusion would it be right to draw if we saw a barometric column standing at considerably over 34 inches ? It might seem plausible to argue that there was no known theoretical limit to the barometric height, from the extreme operation of ordinary meteorological causes, and that it was unphilosophical to go outside this known *vera causa.* But a meteorologist would rightly be unmoved by these arguments. He would know that if this point of view were justified, barometric height of 33 inches would be fairly common and barometric heights of 32 inches very common ; and since this is not the case he would be driven to seek an explanation of the prodigy in some entirely different direction.* For similar reasons we must class auroral displays as quite a distinct set of phenomena.

As we approach the auroral zone, the occurrence of auroral displays becomes more frequent. These, however, at Victoria and at Lerwick are commonly low down in the north, and do not then appreciably affect readings taken 45° up. This is well established by readings at Lerwick taken on many nights at 45° up to north and south. Practically, when the polar aurora is recorded low down in the north, the intensities at 45° up are found to be normal, and the same north and south. Little difficulty was found in picking out those cases where the polar aurora had an important effect. Thus, if we combine the observations recorded at Lerwick and at Victoria from September, 1925, to April, 1927, into a single set, rejecting those which are markedly affected by an auroral display, we have

Number of observations	178
Mean value	+0·6
Standard deviation	1·4
Range.............................	−4·5 to +4·0

Of these values the lowest shows 3·6 times the standard deviation from the mean, and the highest 2·4 times. These conform fairly well to the ordinary rule of three times the standard.†

The values markedly influenced by an auroral display were 11 in number. They were as follows : +12·6, +16·7, +16·7, +6·7, +17·7, +17·7, +13·5, +9·5, +17·5, +17·5, +5·5. This is the complete list, subject to the remark that only one observation is included for any one night, giving preference to that in the north, which, in the case of a difference, is always the brighter.

---

* This illustration is not altogether a fanciful one. Prof. Rowland, of Baltimore, was (I have been told) once confronted with this actual case. The view which he took was that the mercury of the barometric column contained a large quantity of sodium in solution.

† In view of the difficulties of measurement, no importance can be attached to the slight excess of range on the low side.

It will be observed that there is only one case, +5·5, where any doubt can arise as to which category is to be chosen.

As is well known, the visual light of an auroral display consists for the most part of the visual green line, and accordingly it is this component (auroral green) which is most strongly reinforced when a bright display is in progress. It is always found, however, that the red and blue components are strengthened as well. The following typical examples may be given :—

Place.	Date.	Red.	Auroral.	Blue.
Victoria	September 20, 1925	+2·6	+13·5	+13·0
Kingston	March 5, 1926	+5·9	+13·0	+13·4
Lerwick	January 13, 1926	+2·1	+16·7	+11·9

These may be compared with the normal values in the absence of a display, given in § 3. The increase of intensity in the blue part of the spectrum is probably attributable in the main to the negative bands of nitrogen.

The increase in the red is less marked. It may be due in part to the known red line of the aurora, but I believe more to an (apparently) continuous background, which may, of course, be an unresolved band spectrum.

§ 7. *Summary.*

The intensity of the light of the sky at night has been observed by collaborating workers at a number of stations scattered over the world. As in previous work, colour filters are used to separate as nearly as may be the region of the green auroral from the red and blue regions of the spectrum on either side of it. These are called red, auroral and blue components respectively.

The broad result is that the intensity in each component has the same general values with the same general range of intensities (four or fivefold) that have previously been found for England.

Much of the variation is irregular, that is to say, it is not related to any definite period. This part of the variation has not been found to be correlated at the different stations. The stations are, however, for the most part widely separated, and further investigation with stations nearer together is required, in order to determine the distances over which local peculiarities extend.

Five successive years' observation in England appears to indicate a definite annual periodicity, which survives the process of averaging the same month (February, *e.g.*) for each year. The amplitude of variation corresponds to an intensity ratio of 1·6. The maximum is in October.

The available evidence (which, however, is insufficient) suggests a similar variation in the southern hemisphere, with the opposite phase, the maximum occurring in April.

In addition to the annual variation the observations in England (which alone extend over a long enough time) indicate a progressive increase from year to year, the mean annual value having increased 1·3 times during a 4-year interval. This increase has occurred during the rise from sunspot minimum to sunspot maximum, and is thought to indicate a periodicity in the 11-years' period of the solar cycle. Observations are being continued.

A statistical analysis of the intensities at stations which are far enough north for the display of polar aurora, confirms the conclusions drawn before from other evidence and shows that the polar aurora is a distinct phenomenon, incapable of being classified in the same scheme of frequency distribution of the luminous intensity.

In conclusion, I must emphasise that much of the work of this paper, though necessarily organised and edited from one centre, would have been impossible without the friendly and generous co-operation of the workers whose names have already been recited. They have given their time and trouble without stint.

As regards the future, it has been borne in upon me that greater accuracy of photometric measurement is a chief requisite. This will be obtained by means of a photoelectric cell. Most of the difficulties have been overcome, and preliminary observations have been in progress for some months past. I have been able to follow the changes of intensity from hour to hour on clear nights. Although (§ 4) no correlation has been found in the irregular variations at distant stations, there must be correlation at stations close together, and it is important to determine limits of distance. By the photoelectric method this should be comparatively quick and easy.

On the Brownian Displacements and Thermal Diffusion of Grains suspended in a Non-Uniform Fluid.

By S. Chapman, M.A., D.Sc., F.R.S., Imperial College of Science and Technology.

(Received February 2, 1928.—Revised February 28, 1928.)

1. In two famous papers* on the Brownian motion of grains suspended in a stationary uniform fluid (liquid or gaseous) Einstein obtained, *inter alia*, the distribution function for the displacements of the grains during any interval $t = 0$ to $t = \tau$ from their positions at time $t = 0$. The object of this paper is to determine the distribution function for the more general case of a non-uniform fluid. The non-uniformity may refer to temperature, composition, or any other property which affects the coefficient of diffusion (D) of the grains in the fluid. The distribution function is given in § 8, where it is shown how its accuracy might be experimentally tested. It contains terms additional to the one given by Einstein for the uniform case; certain of these are definitely determined, but another important term contains a coefficient that cannot be evaluated by considerations of the kind used in this paper (which depend purely on the conservation of the number of grains), but requires more detailed physical analysis; it is surmised that this coefficient vanishes in the case of Brownian particles which are large compared with the mean free path of the surrounding molecules.

The main results of the paper refer to the rate of diffusion of the grains due to the non-uniformity of the fluid (§ 9), and to the equilibrium distribution of the grains (§§ 10, 11). It is found that if their density is the same as that of the fluid, so that there is no tendency for them to settle in the lower strata, their steady distribution when the fluid is non-uniform is such that the concentration n (or number per unit volume of the fluid) is inversely proportional to D; a solution is also given for the case when the densities are not equal.

The most obvious applications of these results are to the case of fluids at non-uniform temperature T (absolute); since D increases with T in the case of ordinary gases, the grains when suspended in a gaseous medium will tend towards the cooler regions. The same is true also for liquids, where, indeed, the effect will be much more pronounced. This is another example of what, in the case of gases, I have called thermal diffusion, a phenomenon first predicted theoretically by Enskog[*] and myself[†] independently, and afterwards verified by Dootson, Ibbs and Elliott and Masson.[‡] A similar effect in liquid solutions has long been known; it was first described by Ludwig[§] in 1856, and later, independently, by Soret[||] in 1879. Many later workers[¶] have investigated the phenomenon experimentally, and various attempts[**] have been made to construct a theory of it, but I know of none which can be considered satisfactory.

---

* 'Ann. Physik,' vol. 17, p. 549 (1905); and vol. 19, p. 371 (1906).

The results of §§ 9-11 of this paper depend on the assumption referred to in the first paragraph, and cannot be regarded as established. Confirmation must therefore await either experiment or further theoretical progress. In an accompanying paper I have shown that the coefficient in question (C, § 7) does *not* vanish, in general, when the Brownian particles are replaced by ordinary molecules in a perfect gas mixture; for very heavy large molecules the concentration-gradient due to temperature exceeds that given by the relation $n\text{D} = \text{constant}$.

The present theory may be applicable to the Ludwig-Soret effect in liquid *solutions* in cases where the solute molecules are large and non-ionised; this is a matter which should be submitted to an experimental test. But such a test would give information rather as to the character of the solute molecules than as to the validity of the present theory. The latter should be experimentally tested by direct investigation of suspended grains both in liquids and gases, as was done in the case of Einstein's theory by Perrin[††] and Millikan[‡‡] and their collaborators.

After this paper was written I became acquainted with an able and interesting investigation by Wereide,[§§] who gave two alternative theories of the Ludwig-Soret effect, applicable in different circumstances; one of his theories assumes that the solute molecules behave like Brownian particles, and in this case his result for the steady distribution is $n \propto \mathrm{D}^{-\frac{1}{2}}$, which differs from that here obtained.

---

[*] 'Phys. Z.,' vol. 12, p. 538 (1911); 'Ann. Physik,' vol 38, p. 742 (1912); 'Dissertation,' Uppsala, 1917; 'Arkiv f. Nat. Astr. o. Fysik,' vol. 16 (1921).

[†] 'Roy. Soc. Proc.,' A, vol. 93, p. 1 (1916); 'Phil. Trans.,' A, vol. 217, p. 157 (1917).

[‡] Chapman and Dootson, 'Phil. Mag.,' vol. 33, p. 268 (1917); Ibbs, 'Roy. Soc. Proc.,' A, vol. 99, p. 385 (1921), vol. 107, p. 470 (1925), 'Proc. Phys. Soc.,' vol. 39, p. 227 (1927); Elliott and Masson, 'Roy. Soc. Proc.,' A, vol. 108, p. 378 (1925).

[§] 'Wien. Ak. Ber.,' vol. 20, p. 539 (1856).

[||] 'Arch. de Genève,' vol. 2, p. 48 (1879), vol. 4, p. 209 (1880); 'Ann. Chim. Phys.,' vol. 22, p. 293 (1881).

[¶] C. C. Tanner, 'Trans. Faraday Soc.,' vol. 23, p. 75 (1927); and many other references there given.

[**] Arrhenius, 'K. Vet. Ak. Förh.,' vol. 2, p. 61 (1894); Porter, 'Trans. Faraday Soc.,' vol. 23, p. 314 (1927).

[††] 'Les Atomes,' and other references there given.

[‡‡] 'The Electron,' and other references there given.

[§§] 'Ann. Physique,' vol. 2, p. 67 (1914).

2. It will be supposed that the grains are small and spherical, and that their mutual distances are large enough, in relation to their own dimensions, for their thermal motions to be independent of one another to a sufficient degree of approximation. The fluid in which they are suspended may be either a liquid or a gas.

Let τ_0 be the minimum interval such that the displacements of a grain in successive intervals τ_0 are independent of one another. In his second paper on the Brownian motion (§ 5), Einstein has shown how τ_0 may be estimated; for example, it is about 10^{-7} second for a grain of 10^{-4} cm. diameter and of unit density, in water at room temperature. Intervals shorter than τ_0 will not be considered, and $\partial f/\partial t$ will denote the limit of $\Delta f/\tau$, where Δf is the change in the function f during an interval $\tau > \tau_0$, as τ decreases to τ_0. Such a definition is, of course, familiar in statistical problems; it is assumed that the limit exists.

The concentration of grains, n, at time t in the neighbourhood of the point whose Cartesian co-ordinates are x, y, z, and whose position-vector is $\mathbf{r}$, relative to an assigned reference frame, will in general be a function of these variables, and will be written $n(\mathbf{r}, t)$ when it is desired to indicate to which point-instant n refers. The fluid in which the grains are suspended is assumed to be at rest, relative to the reference frame, except in the immediate vicinity of the grains.

Let $\boldsymbol{\rho}$ or ξ, η, ζ, denote the displacement of a typical grain in the interval t to $t + \tau$, starting from the point $\mathbf{r}$. Let us denote by $d\mathbf{r}$ a volume-element of any form surrounding the point $\mathbf{r}$, and by $d\boldsymbol{\rho}$ a volume element surrounding the point $\mathbf{r} + \boldsymbol{\rho}$; $d\mathbf{r}$ or $d\boldsymbol{\rho}$ is to be carefully distinguished from $\mathbf{dr}$ or $\mathbf{d}\boldsymbol{\rho}$, which in ordinary vector usage denotes a vector increment of $\mathbf{r}$ or $\boldsymbol{\rho}$; if the volume-element $d\mathbf{r}$ or $d\boldsymbol{\rho}$ is a parallelepiped with its edges parallel to the co-ordinate axes, it will be expressible as $dx\,dy\,dz$ or $d\xi\,d\eta\,d\zeta$, but it is often convenient to use spherical polar or other co-ordinates; if spherical polar co-ordinates are used, $d\mathbf{r}$ would be expressed as $r^2 \sin\theta\,dr\,d\theta\,d\phi$. The volume-element will always be regarded as positive; if the co-ordinates in which it is expressed are altered, the order of the limits will be supposed altered, if necessary, in order that $d\mathbf{r}$ or $d\boldsymbol{\rho}$ may remain positive in its new form.

After an interval τ the $n d\mathbf{r}$ grains in $d\mathbf{r}$ at time t will be displaced by different amounts and in different directions. The displacements can be represented by a set of vectors OP all starting from the same point O in an auxiliary space, and the distribution of the displacements can be specified by the distribution of the ends P of the vectors. Let the number of these end-points lying within a volume-element $d\boldsymbol{\rho}$ (of any form) surrounding the point P such that OP $= \boldsymbol{\rho}$ be denoted by

$$(1) \qquad n(\mathbf{r}, t)\, \phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)\, d\mathbf{r}\, d\boldsymbol{\rho}\,;$$

this is clearly the number of grains in $d\mathbf{r}$ which after time τ have been displaced approximately (within limits defined by $d\boldsymbol{\rho}$) by the amount $\boldsymbol{\rho}$. By summing up the number of end-points in all volume-elements $d\boldsymbol{\rho}$ in the auxiliary space we must obtain $n d\mathbf{r}$, the total number of grains under consideration; hence, representing the summation in integral notation, we have

$$(2) \qquad \int \phi\, d\boldsymbol{\rho} = 1,$$

where, as always in this paper when limits are not specially indicated, the integration extends over the whole range of the vector variable (in this case $\boldsymbol{\rho}$). The function ϕ, which is essentially positive, will be called the displacement-distribution function: frequently it will be denoted simply by $\phi(\boldsymbol{\rho}, \tau)$, the other variables not being indicated.

It is convenient to write

$$(3) \qquad \int f(\boldsymbol{\rho})\, \phi(\boldsymbol{\rho}, \tau)\, d\boldsymbol{\rho} = \overline{f},$$

where f denotes any function of $\boldsymbol{\rho}$; $\overline{f}$ thus denotes the mean value of f for the set of grains considered. In general $\overline{f}$ will depend on τ and $\mathbf{r}$, t, because of the varying distribution of the displacements at different times and in different parts of the fluid; when it is desired to indicate this, $\overline{f}$ will be written $\overline{f}(\tau, \mathbf{r}, t)$. For example, when $f(\boldsymbol{\rho}) \equiv \xi^a \eta^b \zeta^c$, the mean value will be denoted by $\overline{\zeta^a \eta^b \zeta^c}(\tau, \mathbf{r}, t)$ or, more briefly, by $\overline{\xi^a \eta^b \zeta^c}$.

When the interval τ is made equal to τ_0, $\phi(\boldsymbol{\rho}, \tau_0, \mathbf{r}, t)$ will be denoted briefly by $\phi_0(\boldsymbol{\rho})$ or ϕ_0, and $\overline{f}$ will likewise be denoted briefly by $\overline{f}_0$.

3. The number of grains in $d\mathbf{r}$ at the point-instant $\mathbf{r}$, $t + \tau$ can be found by considering how the distribution of grains throughout the fluid at the time t will contribute to this volume element after the interval τ, all the original grains having then left the element $d\mathbf{r}$. This gives the equation

$$(4) \qquad n(\mathbf{r}, t + \tau)\, d\mathbf{r} = d\mathbf{r} \int n(\mathbf{r} - \boldsymbol{\rho}, t)\, \phi(\boldsymbol{\rho}, \tau, \mathbf{r} - \boldsymbol{\rho}, t)\, d\boldsymbol{\rho}.$$

The integrand, regarded as a function of the vector-variable $\mathbf{r} - \boldsymbol{\rho}$, or the three scalar variables $x - \xi$, $y - \eta$, $z - \zeta$, can be expanded by Taylor's theorem in powers of ξ, η, ζ; as such expansions will be used several times, and are very cumbrous to write out in full, it is convenient to use vector notation, in the form

$$\mathrm{F}(\mathbf{r}-\boldsymbol{\rho}) = \sum_{s=0}^{\infty} \frac{1}{s!}\left(-\boldsymbol{\rho}, \frac{\partial}{\partial \mathbf{r}}\right)^{s} \mathrm{F}(\mathbf{r}),$$

where

$$\left(\boldsymbol{\rho}, \frac{\partial}{\partial \mathbf{r}}\right) \equiv \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \zeta\frac{\partial}{\partial z}, \tag{5}$$

and powers of this differential operator are to be expanded by the multinomial theorem.

Hence (4) may be written (after division by $d\mathbf{r}$)

$$n(\mathbf{r}, t+\tau) = \int \sum_{s=0}^{\infty} \frac{1}{s!}\left(-\boldsymbol{\rho}, \frac{\partial}{\partial \mathbf{r}}\right)^{s} \{n(\mathbf{r}, t)\, \phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)\}\, d\boldsymbol{\rho}.$$

A typical term in the expanded integral on the right is

$$\frac{1}{s!}(-1)^{a+b+c}\, {}_{s}\mathrm{C}_{a,b,c} \left(\frac{\partial}{\partial x}\right)^{a} \left(\frac{\partial}{\partial y}\right)^{b} \left(\frac{\partial}{\partial z}\right)^{c} \left\{n(\mathbf{r}, t) \int \xi^{a}\eta^{b}\zeta^{c}\phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)\, d\boldsymbol{\rho}\right\}$$
$$= \frac{1}{s!}(-1)^{a+b+c}\, {}_{s}\mathrm{C}_{a,b,c} \left(\frac{\partial}{\partial x}\right)^{a} \left(\frac{\partial}{\partial y}\right)^{b} \left(\frac{\partial}{\partial z}\right)^{c} \{n(\mathbf{r}, t)\, \overline{\xi^{a}\eta^{b}\zeta^{c}}(\tau, \mathbf{r}, t)\},$$

where $a+b+c = s$, and ${}_{s}\mathrm{C}_{a,b,c}$ is the multinomial coefficient of $\mathrm{X}^{a}\mathrm{Y}^{b}\mathrm{Z}^{c}$ in the expansion of $(\mathrm{X}+\mathrm{Y}+\mathrm{Z})^{s}$.

A convenient vector notation for the new form of (4) thus obtained is

$$n(\mathbf{r}, t+\tau) = \sum_{s=0}^{\infty} \frac{1}{s!}\overline{\left(\frac{\partial}{\partial \mathbf{r}}, -\boldsymbol{\rho}\right)^{s}}\, n, \tag{6}$$

the independent variables on the right being understood to be τ, $\mathbf{r}$, t. The term in (6) corresponding to $s = 0$ is simply $n(\mathbf{r}, t)$; on transferring this to the left, dividing by τ, and letting τ take the value τ_0, we obtain

$$\begin{aligned}\frac{\partial n(\mathbf{r}, t)}{\partial t} &= \frac{1}{\tau_0}\sum_{s=1}^{\infty} \frac{1}{s!}\overline{\left(\frac{\partial}{\partial \mathbf{r}}, -n\boldsymbol{\rho}_0\right)^{s}} \\ &= -\frac{1}{\tau_0}\left\{\frac{\partial}{\partial x}(n\overline{\xi}_0) + \frac{\partial}{\partial y}(n\overline{\eta}_0) + \frac{\partial}{\partial z}(n\overline{\zeta}_0)\right\} \\ &\qquad + \frac{1}{2!\,\tau_0}\left\{\frac{\partial^2}{\partial x^2}(n\overline{\xi_0^2}) + 2\frac{\partial^2}{\partial y\, \partial z}(n\overline{\eta_0\zeta_0}) + \dots\right\} + \dots,\end{aligned} \tag{7}$$

where on the right n and the mean values refer to $\mathbf{r}$, t (the suffix 0 referring—*cf.* § 2—to the fact that they relate to an interval τ_0).

It will appear later that, on the assumption mentioned in § 1 and described in § 8, the only terms on the right-hand side of (7) which have to be considered when the fluid is non-uniform and the grains are subject to no external forces are those in the last line, involving $\overline{\xi_0^2}, \overline{\eta_0^2}, \overline{\zeta_0^2}$: and that the phenomenon of Brownian thermal diffusion theoretically predicted in this paper depends essentially on the fact that these mean values are under the operation of the *double* differentiations with respect to x, y, z.

4. In a somewhat similar way we may obtain a differential equation which indicates how ϕ must depend on $\boldsymbol{\rho}$ and τ. The number of grains which, originally in $d\mathbf{r}$ at $\mathbf{r}, t$, are in the element $d\boldsymbol{\rho}'$ about the point $\mathbf{r}+\boldsymbol{\rho}'$ at time $t+\tau$, is

$$n(\mathbf{r},t)\,\phi(\boldsymbol{\rho}',\tau,\mathbf{r},t)\,d\boldsymbol{\rho}'\,d\mathbf{r}.$$

Similarly the number of those which are in the element $d\boldsymbol{\rho}$ at the point $\mathbf{r}+\boldsymbol{\rho}$ at the time $t+\tau+\tau_0$ is

$$(8)\qquad n(\mathbf{r},t)\,\phi(\boldsymbol{\rho},\tau+\tau_0,\mathbf{r},t)\,d\boldsymbol{\rho}\,d\mathbf{r}.$$

But the latter number can be calculated by considering how many of the grains originally in $d\mathbf{r}$ at $\mathbf{r}, t$ move from their positions at time $t+\tau$ into the element $d\boldsymbol{\rho}$ at $\mathbf{r}+\boldsymbol{\rho}$ in the succeeding interval τ_0. The total number of grains in $d\boldsymbol{\rho}'$ at $\mathbf{r}+\boldsymbol{\rho}'$ at $t+\tau$ is $n(\mathbf{r}+\boldsymbol{\rho}', t+\tau)\,d\boldsymbol{\rho}'$, and of these the number which were in $d\mathbf{r}$ at $\mathbf{r}, t$ is

$$n(\mathbf{r},t)\;\phi(\boldsymbol{\rho}',\tau,\mathbf{r},t)\,d\boldsymbol{\rho}'\,d\mathbf{r}.$$

The motion of the grains in $d\boldsymbol{\rho}'$ during the further interval τ_0 is, by hypothesis, independent of their motions in the previous interval τ. Since the number of them which at $t+\tau+\tau_0$ are in $d\boldsymbol{\rho}$ at $\mathbf{r}+\boldsymbol{\rho}$ is

$$n(\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,\phi(\boldsymbol{\rho}-\boldsymbol{\rho}',\tau_0,\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,d\boldsymbol{\rho}'\,d\mathbf{r},$$

the number which were originally in $d\mathbf{r}$ is

$$\frac{n(\mathbf{r},t)\,\phi(\boldsymbol{\rho}',\tau,\mathbf{r},t)\,d\boldsymbol{\rho}'\,d\mathbf{r}}{n(\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,d\boldsymbol{\rho}'}\,n(\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,\phi(\boldsymbol{\rho}-\boldsymbol{\rho}',\tau_0,\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,d\boldsymbol{\rho}'\,d\boldsymbol{\rho}$$

or

$$(9)\qquad n(\mathbf{r},t)\,\phi(\boldsymbol{\rho}',\tau,\mathbf{r},t)\,\phi(\boldsymbol{\rho}-\boldsymbol{\rho}',\tau_0,\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,d\boldsymbol{\rho}'\,d\boldsymbol{\rho}\,d\mathbf{r}.$$

To get the total number of grains in $d\boldsymbol{\rho}$ at $\mathbf{r}+\boldsymbol{\rho}$, $t+\tau+\tau_0$ which were originally in $d\mathbf{r}$ at $\mathbf{r}, t$, (9) must be integrated over the whole range of $\boldsymbol{\rho}'$. Equating the result to (8), we get

$$\begin{aligned}(10)\qquad \phi(\boldsymbol{\rho},\tau+\tau_0,\mathbf{r},t) &= \int\phi(\boldsymbol{\rho}',\tau,\mathbf{r},t)\,\phi(\boldsymbol{\rho}-\boldsymbol{\rho}',\tau_0,\mathbf{r}+\boldsymbol{\rho}',t+\tau)\,d\boldsymbol{\rho}'\\ &= \int\phi(\boldsymbol{\rho}-\boldsymbol{\rho}'',\tau,\mathbf{r},t)\,\phi(\boldsymbol{\rho}'',\tau_0,\mathbf{r}+\boldsymbol{\rho}-\boldsymbol{\rho}'',t+\tau)\,d\boldsymbol{\rho}'',\end{aligned}$$

where $\boldsymbol{\rho}'' = \boldsymbol{\rho} - \boldsymbol{\rho}'$; the volume element $d\boldsymbol{\rho}''$, like the rest of the integrand in both integrals, is positive (*cf.* § 2).

Expanding the integrand by Taylor's theorem as follows,

$$\sum_{s=0}^{\infty} \frac{1}{s\,!}\left(-\boldsymbol{\rho}'', \frac{\partial}{\partial \boldsymbol{\rho}}\right)^{s} \{\phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)\, \phi(\boldsymbol{\rho}'', \tau_0, \mathbf{r}+\boldsymbol{\rho}, t+\tau)\}, \tag{11}$$

on integration over the whole range of $\boldsymbol{\rho}''$ we get

$$\phi(\boldsymbol{\rho}, \tau+\tau_0, \mathbf{r}, t) = \sum_{s=0}^{\infty} \frac{1}{s\,!}\left(\frac{\partial}{\partial \boldsymbol{\rho}}, -\overline{\boldsymbol{\rho}''}\right)^{s} \phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t), \tag{12}$$

using the same symbolism as in (6); *where on the right the mean values refer to the point-instant* $\mathbf{r}+\boldsymbol{\rho}, t+\tau$ and the interval τ_0, and are therefore functions of these variables.

On transferring the first term on the right to the other side of the equation and dividing by τ_0, we get

$$\begin{aligned}\frac{\partial\, \phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)}{\partial \tau} = -\frac{1}{\tau_0}\left\{\frac{\partial}{\partial \xi}(\bar{\xi}_0 \phi) + \frac{\partial}{\partial \eta}(\bar{\eta}_0\, \phi) + \frac{\partial}{\partial \zeta}(\bar{\zeta}_0, \phi)\right\} \\ + \frac{1}{2\,!\,\tau_0}\left\{\frac{\partial^2}{\partial \xi^2}(\overline{\xi_0^2}\, \phi) + 2\frac{\partial^2}{d\eta\, \partial \zeta}(\overline{\eta_0 \zeta_0} \phi) + \dots\ \right\} + \dots,\end{aligned} \tag{13}$$

where the mean values refer to the point-instant $\mathbf{r}+\boldsymbol{\rho}, t+\tau$, while ϕ has the same variables as on the left of the equation.

5. The mean values in (13) can be expressed in terms of the corresponding mean values at $\mathbf{r}, t$ and the space-time gradients of $\phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)$, by expanding $\phi(\boldsymbol{\rho}'', \tau_0, \mathbf{r}+\boldsymbol{\rho}, t+\tau)$ in (11) in powers of $\boldsymbol{\rho}$ and τ. It is convenient to treat $\boldsymbol{\rho}, \tau$ as together forming a four-dimensional vector $\boldsymbol{\lambda}$, and $\mathbf{r}, t$ as forming a four-dimensional vector $\mathbf{l}$. Then

$$\phi(\boldsymbol{\rho}'', \tau_0, \mathbf{r}+\boldsymbol{\rho}, t+\tau) = \sum_{q=0}^{\infty} \frac{1}{q\,!}\left(\boldsymbol{\lambda}, \frac{\partial}{\partial \mathbf{l}}\right)^{q} \phi(\boldsymbol{\rho}'', \tau_0, \mathbf{r}, t), \tag{14}$$

where

$$\left(\boldsymbol{\lambda}, \frac{\partial}{\partial \mathbf{l}}\right) \equiv \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \zeta\frac{\partial}{\partial z} + \tau\frac{\partial}{\partial t}. \tag{15}$$

Consequently

$$\overline{\xi^a \eta^b \zeta^c}(\tau_0, \mathbf{r}+\boldsymbol{\rho}, t+\tau) = \sum_{q=0}^{\infty} \frac{1}{q\,!}\left(\boldsymbol{\lambda}, \frac{\partial}{\partial \mathbf{l}}\right)^{q} \overline{\xi^a \eta^b \zeta^c}(\mathbf{r}, t). \tag{16}$$

After substitution in (13) from (16), the former becomes a differentio-integral equation for ϕ, since the mean values are integrals involving ϕ. The equation is to some extent analogous to Boltzmann's equation for the velocity-distribution function in the kinetic theory of gases.

6. The equation (13) takes a specially simple form when the fluid in which the grains are suspended is uniform, and no external force, such as gravity, renders their motions unsymmetrical. In this case all the space-time gradients of mean values, such as occur on the right-hand side of (16), are zero, *i.e.*, $\overline{\xi^a\eta^b\zeta^c}$ is independent of $\mathbf{r}$, t; moreover, it is zero unless a, b, c are all even, and it is unaltered by permuting the indices of ξ, η, ζ. Thus (13) reduces to

$$\frac{\partial\phi(\boldsymbol{\rho},\tau)}{\partial\tau}=\sum_{s=1}^{\infty}\frac{1}{(2s)!\,\tau_0}\overline{\left(\boldsymbol{\rho}_0,\frac{\partial}{\partial\boldsymbol{\rho}}\right)^{2s}}\phi(\boldsymbol{\rho},\tau), \tag{17}$$

in which the variables $\mathbf{r}$, t no longer appear.

Einstein obtained a solution of this special form of the equation (13) by ignoring all terms on the right after the first ($s=1$); (17) then becomes

$$\begin{aligned}\frac{\partial\phi}{\partial\tau}&=\frac{1}{2\tau_0}\left(\overline{\xi_0^2}\frac{\partial^2}{\partial\xi^2}+\overline{\eta_0^2}\frac{\partial^2}{\partial\eta^2}+\overline{\zeta_0^2}\frac{\partial^2}{\partial\zeta^2}\right)\phi\\&=\tfrac{1}{6}\frac{\overline{\rho_0^2}}{\tau_0}\nabla^2\phi,\end{aligned} \tag{18}$$

where

$$\nabla^2\equiv\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}+\frac{\partial^2}{\partial\zeta^2}. \tag{19}$$

We have here used the result, due to the symmetry of the displacements,

$$\overline{\xi_0^2}=\overline{\eta_0^2}=\overline{\zeta_0^2}=\tfrac{1}{3}\overline{(\xi_0^2+\eta_0^2+\zeta_0^2)}=\tfrac{1}{3}\overline{\rho_0^2}=\tfrac{1}{3}\int\rho^2\phi(\boldsymbol{\rho},\tau_0)\,d\boldsymbol{\rho}; \tag{20}$$

the distinction between the vector $\boldsymbol{\rho}$ and its magnitude ρ should be noted.

Writing

$$D=\tfrac{1}{6}(\overline{\rho_0^2}/\tau_0), \tag{21}$$

(18) becomes

$$\frac{\partial\phi}{\partial\tau}=D\nabla^2\phi. \tag{22}$$

The solution of this equation, possessing symmetry with respect to ξ, η, ζ, finite at $\boldsymbol{\rho}=0$, zero at $\boldsymbol{\rho}=\infty$, and satisfying (2), is

$$\phi=\frac{1}{(4\pi D\tau)^{3/2}}e^{-\rho^2/4D\tau}. \tag{23}$$

On substituting from (23) in (3) and writing $d\boldsymbol{\rho}=d\xi\,d\eta\,d\zeta$, it is evident that mean values of odd functions of ξ, η, ζ are zero, and it is readily verified that $\overline{\zeta^{2a}\eta^{2b}\zeta^{2c}}/\tau$ is a purely numerical multiple of

$$(\mathrm{D}\tau)^{a+b+c}/\tau. \tag{24}$$

When $a + b + c = 1$ the result is independent of τ, and, in fact,

$$\frac{\overline{\xi^2}}{\tau} = \frac{\overline{\xi_0{}^2}}{\tau_0} = \tfrac{1}{3}\frac{\overline{\rho_0{}^2}}{\tau_0} \doteq 2\mathrm{D}, \tag{25}$$

as is required by (21). When $a + b + c > 1$, $\overline{\xi_0{}^{2a}\,\eta_0{}^{2b}\,\zeta_0{}^{2c}}/\tau_0$ is a numerical multiple of $\tau_0{}^{a+b+c-1}$, which is very small if, as we shall throughout suppose, τ_0 is very small. This justifies the neglect of all terms save the first, on the right of (17).

7. We now proceed to consider (13) in the case when the grains are subject to external forces $\mathbf{F}$ (depending only on their position), such as gravitational forces due to a difference between their density and that of the fluid, or electrical forces; and when the fluid is non-uniform in regard to some scalar property α. The displacement-distribution function ϕ will now be a function of $\mathbf{r}$, t as well as of $\boldsymbol{\rho}$, τ, but we suppose that $\mathbf{F}$ and the non-uniformity of the fluid are small enough for their effects upon ϕ to be represented by first-order corrections, involving $\mathbf{F}$ or its components X, Y, Z, and the gradient of α, which it is convenient to denote by $\partial\alpha/\partial\mathbf{r}$, having components

$$\frac{\partial\alpha}{\partial x},\quad \frac{\partial\alpha}{\partial y},\quad \frac{\partial\alpha}{\partial z}\,.$$

Since the expression for ϕ must be invariant with respect to changes of orientation of the orthogonal Cartesian axes, we may write

$$\phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t) = \Phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)\left\{1 + \left(\frac{\partial\alpha}{\partial\mathbf{r}}, \boldsymbol{\rho}\right)\chi(\rho^2, \tau, \mathbf{r}, t) + (\mathbf{F}, \boldsymbol{\rho})\,\psi(\rho^2, \tau, \mathbf{r}, t)\right\}, \tag{26}$$

where

$$\Phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t) \equiv \frac{1}{(4\pi\mathrm{D}\tau)^{3/2}}\,e^{-\rho^2/4\mathrm{D}\tau}, \tag{27}$$

and

$$\mathrm{D} \equiv \tfrac{1}{6}\frac{\overline{\rho_0{}^2}}{\tau_0} \tag{28}$$

at $\mathbf{r}$, t; thus D in (27) is now a function of $\mathbf{r}$, t. Moreover, as (26) indicates,

χ and ψ are functions which depend only on the magnitude of $\boldsymbol{\rho}$ and not on its direction (this follows from the necessity for invariance of the whole expression with respect to orientation of the axes).

If the fluid is non-uniform with respect to more than one (scalar) property α, an additional term for each such property must be inserted in (26), each with its own distinct co-factor χ, which may be different for different α's.

When $\mathbf{F}$ and $\partial\alpha/\partial\mathbf{r}$ are zero, ϕ reduces to Φ, as in § 6. It may also be noted that (28) is consistent with (26) and (27), since

$$\overline{\rho_0^2} = \int \rho^2 \phi(\boldsymbol{\rho}, \tau_0, \mathbf{r}, t)\, d\boldsymbol{\rho} = \int \rho^2 \Phi_0\, d\boldsymbol{\rho} + \left(\frac{\partial\alpha}{\partial\mathbf{r}}, \int \boldsymbol{\rho}\rho^2\chi_0\Phi_0\, d\boldsymbol{\rho}\right) + \left(\mathbf{F}, \int \boldsymbol{\rho}\rho^2\psi_0\Phi_0\, d\boldsymbol{\rho}\right)$$

$$= \int \rho^2 \Phi_0\, d\boldsymbol{\rho} = 6\mathrm{D}\tau_0,$$

since the integrals involving χ_0 and ψ_0 are of odd degree in $\boldsymbol{\rho}$; in these expressions, as elsewhere in the paper, the suffix $_0$ indicates that $\tau = \tau_0$.

In consequence of the additional terms in ϕ, mean values such as $\overline{\xi^a\eta^b\zeta^c}$ will now depend on $\mathbf{r}$, t, and they will not vanish, in general, even when (not more than) one of a, b, c is odd. We have

$$\frac{1}{\tau_0}\overline{\xi_0{}^a\eta_0{}^b\zeta_0{}^c} = \frac{1}{\tau_0}\overline{[\xi_0{}^a\eta_0{}^b\zeta_0{}^c]} + \frac{1}{\tau_0}\left(\frac{\partial\alpha}{\partial\mathbf{r}}, \int \boldsymbol{\rho}\chi_0\xi^a\eta^b\zeta^c\Phi_0\, d\boldsymbol{\rho}\right)$$

$$+ \frac{1}{\tau_0}\left(\mathbf{F}, \int \boldsymbol{\rho}\psi_0\xi^a\eta^b\zeta^c\Phi_0\, d\boldsymbol{\rho}\right),$$

in which, on the right, the first term denotes the mean value corresponding to the case of a uniform fluid having the same value of D as the actual fluid at $\mathbf{r}$, t, and to the absence of external forces on the grains; it is therefore zero when a, b, c are not all even, and when they are it is a multiple of $\tau_0{}^{\frac{1}{2}(a+b+c)-1}$ and consequently negligible except when two of a, b, c are zero and the third is 2. Hence, except in this case, when its value is 2D (*cf.* (25)), the first term is negligible. In the other two terms the integrals and integrands are vectors, on account of the factor $\boldsymbol{\rho}$; each component can be considered separately. When a, b, c are all even the integrals vanish, the integrands being of odd degrees; the same applies when more than one of a, b, c is odd. When only one of them is odd, the corresponding component of $\boldsymbol{\rho}$ renders that component of the integrand even, and non-zero. But even in this case the integrals steadily decrease, since every rise by 2 of the combined degrees $a + b + c$ introduces a new factor τ_0 into the result. Hence all these terms are negligible compared with that for which $a+b+c=1$; and in this case, taking $a=1$, $b=0$, $c=0$ as typical,

$$\bar{\xi}/\tau = \frac{1}{\tau}\left(\frac{\partial\alpha}{\partial\mathbf{r}}, \int\boldsymbol{\rho}\xi\chi\Phi\,d\boldsymbol{\rho}\right) + \frac{1}{\tau}\left(\mathbf{F}, \int\boldsymbol{\rho}\xi\psi\Phi\,d\boldsymbol{\rho}\right) \tag{29}$$

$$= \frac{1}{\tau}\frac{\partial\alpha}{\partial x}\int\xi^2\chi\Phi\,d\boldsymbol{\rho} + \frac{1}{\tau}\mathrm{X}\int\xi^2\psi\Phi\,d\boldsymbol{\rho}$$

$$= \frac{1}{\tau}\frac{\partial\alpha}{\partial x}\overline{[\xi^2\chi]} + \frac{1}{\tau}\mathrm{X}\overline{[\xi^2\psi]}$$

$$= \frac{1}{3\tau}\left\{\overline{[\rho^2\chi]}\frac{\partial\alpha}{\partial x} + \overline{[\rho^2\psi]}\mathrm{X}\right\},$$

where again the square brackets denote reference to mean values for the case of a uniform fluid and no external forces on the grains.

Hence

$$\bar{\xi}_0/\tau_0 = \mathrm{C}\frac{\partial\alpha}{\partial x} + \mathrm{C}'\mathrm{X}, \tag{30}$$

where

$$\mathrm{C} = \frac{1}{3\tau_0}\overline{[\rho_0{}^2\chi_0]}, \qquad \mathrm{C}' = \frac{1}{3\tau_0}\overline{[\rho_0{}^2\psi_0]}, \tag{31}$$

C, C′ being mean values which may be of different magnitude at different points in the fluid.

On substituting the mean values thus obtained into (13), and ignoring all negligible terms, we get

$$\frac{\partial\phi(\boldsymbol{\rho}, \tau, \mathbf{r}, t)}{\partial\tau} = -\left(\frac{\partial\alpha}{\partial\mathbf{r}}, \frac{\partial\mathrm{C}\phi}{\partial\boldsymbol{\rho}}\right) - \left(\mathbf{F}, \frac{\partial\mathrm{C}'\phi}{\partial\boldsymbol{\rho}}\right) + \nabla^2(\mathrm{D}\phi), \tag{32}$$

where on the right the mean values C, C′ and D, but not ϕ, refer to the point-instant $\mathbf{r}+\boldsymbol{\rho}$, $t+\tau$. But since $\partial\alpha/\partial\mathbf{r}$ and $\mathbf{F}$ are by hypothesis small, in the first two terms on the right we may give C and C′ their values at $\mathbf{r}$, t and treat them as independent of $\boldsymbol{\rho}$, and we may substitute Φ for ϕ; in the other terms we must substitute for ϕ from (26). In the third term on the right we must take into account the small difference between D at $\mathbf{r}+\boldsymbol{\rho}$, $t+\tau$, and at $\mathbf{r}$, t; the change in t will not matter if, as we shall suppose, the properties of the fluid do not vary with the time; but we must write

$$\mathrm{D} + (\boldsymbol{\rho}, \partial\mathrm{D}/\partial\mathbf{r})$$

or

$$\mathrm{D} + \left(\boldsymbol{\rho}, \frac{d\mathrm{D}}{d\alpha}\frac{\partial\alpha}{\partial\mathbf{r}}\right)$$

in place of D in that term.

Making these substitutions, and ignoring second-order quantities, (32) becomes

$$\frac{\partial\Phi}{\partial\tau}+\boldsymbol{\rho},\left(\frac{\partial\alpha}{\partial\mathbf{r}}\frac{\partial\chi\Phi}{\partial\tau}\right)+(\mathbf{F},\boldsymbol{\rho})\frac{\partial\psi\Phi}{\partial\tau}=-\mathrm{C}\left(\frac{\partial\alpha}{\partial\mathbf{r}},\frac{\partial\Phi}{\partial\boldsymbol{\rho}}\right)-\mathrm{C}'\left(\mathbf{F},\frac{\partial\Phi}{\partial\boldsymbol{\rho}}\right)$$

$$+\left\{\mathrm{D}\nabla^2\Phi+\mathrm{D}\left(\frac{\partial\alpha}{\partial\mathbf{r}},\nabla^2\boldsymbol{\rho}\chi\Phi\right)+\mathrm{D}\,(\mathbf{F},\nabla^2\boldsymbol{\rho}\psi\Phi)+\frac{d\mathrm{D}}{d\alpha}\left(\frac{\partial\alpha}{\partial\mathbf{r}},\nabla^2\boldsymbol{\rho}\Phi\right)\right\}.$$

All the terms here are small, of the first order, except $\partial\Phi/\partial\tau$ and $\mathrm{D}\nabla^2\Phi$, which cancel by (22). The reduction of the remainder of the equation is a lengthy process, in which use is made of the following (and analogous) results:—

$$\frac{\partial\Phi}{\partial\boldsymbol{\rho}}=-\frac{1}{2\mathrm{D}\tau}\boldsymbol{\rho}\Phi,$$

$$\frac{\partial\Phi}{\partial\tau}=\mathrm{D}\nabla^2\Phi=-\frac{1}{\tau}\left(\frac{3}{2}-\frac{\rho^2}{4\mathrm{D}\tau}\right)\Phi.$$

$$\nabla^2(\boldsymbol{\rho}\chi\Phi)=\boldsymbol{\rho}\nabla^2\chi\Phi+2\frac{\partial\chi\Phi}{\partial\boldsymbol{\rho}}.$$

$$\nabla^2(\chi\Phi)=\chi\nabla^2\Phi+2\left(\frac{\partial\chi}{\partial\boldsymbol{\rho}},\frac{\partial\Phi}{\partial\boldsymbol{\rho}}\right)+\Phi\nabla^2\chi,$$

$$\frac{\partial\chi}{\partial\boldsymbol{\rho}}=\frac{\partial\chi}{\partial\rho^2}\frac{\partial\rho^2}{\partial\boldsymbol{\rho}}=2\boldsymbol{\rho}\frac{\partial\chi}{\partial\rho^2},$$

$$\nabla^2\chi=\frac{1}{\rho^2}\frac{\partial}{\partial\rho}\left(\rho^2\frac{\partial\chi}{\partial\rho}\right)=\frac{2}{\rho^2}\frac{\partial}{\partial\rho}\left(\rho^3\frac{\partial\chi}{\partial\rho^2}\right)=4\rho^2\frac{\partial^2\chi}{\partial(\rho^2)^2}+6\frac{\partial\chi}{\partial\rho^2},$$

$$\nabla^2(\boldsymbol{\rho}\Phi)=\boldsymbol{\rho}\nabla^2\Phi+2\frac{\partial\Phi}{\partial\boldsymbol{\rho}}.$$

On making these substitutions, multiplying throughout by τ/Φ, and changing the independent variables in χ and ψ from ρ^2, τ to θ, τ, where

$$\theta=\frac{\rho^2}{4\mathrm{D}\tau},$$

the preceding equation for ϕ becomes

$$(33)\quad\left(\frac{\partial\alpha}{\partial\mathbf{r}},\boldsymbol{\rho}\right)\left\{\theta\chi''+\left(\frac{5}{2}-\theta\right)\chi'-\tau\frac{\partial\chi}{\partial\tau}-\chi+\frac{\mathrm{C}}{2\mathrm{D}}-\left(\frac{5}{2}-\theta\right)\frac{d\log\mathrm{D}}{d\alpha}\right\}$$

$$+(\mathbf{F},\boldsymbol{\rho})\left\{\theta\psi''+\left(\frac{5}{2}-\theta\right)\psi'-\tau\frac{\partial\psi}{\partial\tau}-\psi+\frac{\mathrm{C}'}{2\mathrm{D}}\right\}=0,$$

where (′) denotes differentiation with respect to θ.

Since $\partial\alpha/\partial\mathbf{r}$ and $\mathbf{F}$ are independent of one another, the expressions in brackets { } must separately vanish; hence

$$(34)\quad\theta\chi''+\left(\frac{5}{2}-\theta\right)\chi'-\tau\frac{\partial\chi}{\partial\tau}-\chi=-\frac{\mathrm{C}}{2\mathrm{D}}+\left(\frac{5}{2}-\theta\right)\frac{d\log\mathrm{D}}{d\alpha},$$

and a similar equation in ψ with the last term absent. Hence it is sufficient to determine χ, when ψ can be obtained by omitting the term depending on $d \log D/d\alpha$, and changing C to C′.

The general integral consists of two parts, functions of θ and τ separately; the part depending on τ only is a multiple of $\log \tau$; its coefficient must be zero, for otherwise $\overline{\xi}/\tau$ would contain a term having $\log \tau$ as a factor, implying a mean velocity of the grains which steadily increases to infinity; this is impossible in a fluid having finite viscosity.

Hence the term $\tau\, \partial\chi/\partial\tau$ in (34) can be omitted; the solution of the remaining (ordinary) differential equation for χ is

$$(35)\quad \chi = \frac{C}{2D} + \left(\tfrac{1}{2}\theta - \frac{5}{4}\right)\frac{d \log D}{d\alpha} + A\left(\theta^{-3/2}e^{\theta}\int_{\sqrt{\theta}}^{\infty} e^{-\sigma^2}\, d\sigma + \frac{1}{\theta}\right) + B\theta^{-3/2}\, e^{\theta},$$

where A, B are arbitrary constants. It may be noted also that, in terms of θ,

$$\Phi = \frac{1}{(4\pi D\tau)^{3/2}}\, e^{-\theta}.$$

I am indebted to Mr. C. H. Kebby for help in regard to the expression of the complementary function (the terms containing the factors A, B) in (35) in the above convenient form; he has also kindly checked the analysis of the remainder of the paper.

It is easily seen that $\overline{\xi}/\tau$ cannot be finite unless

$$B = 0.$$

To determine A, we calculate the value of C (*cf.* (31)) by substitution of the above form of χ into the integral giving the mean value. Thus

$$\begin{aligned} C &= \frac{1}{3\tau_0}\overline{[\rho_0^2\chi_0]} = \frac{1}{3\tau_0}\int \rho^2\chi_0\,\Phi_0\, dp = \frac{2}{3\pi^{1/2}\tau_0}\int_0^\infty \rho^2\chi_0\theta_0^{1/2}e^{-\theta_0}d\theta_0 \\ &= \frac{8D}{3\pi^{1/2}}\int_0^\infty \theta_0^{3/2}e^{-\theta_0}\left\{\frac{C}{2D} + \left(\tfrac{1}{2}\theta_0 - \frac{5}{4}\right)\frac{d \log D}{d\alpha}\right. \\ &\qquad\qquad \left. + A\left(\theta_0^{-3/2}e^{-\theta_0}\int_{\sqrt{\theta_0}}^{\infty} e^{-\sigma^2}d\sigma + \frac{1}{\theta_0}\right)\right\} d\theta_0 \\ &= C + A, \end{aligned}$$

since

$$\int_0^\infty x^{5/2}e^{-x}\, dx = \frac{15}{8}\pi^{1/2},\quad \int_0^\infty x^{3/2}e^{-x}\, dx = \frac{3}{4}\pi^{1/2},\quad \int_0^\infty x^{1/2}e^{-x}\, dx = \tfrac{1}{2}\pi^{1/2},$$

and $$\int_0^\infty \int_{\surd x}^\infty e^{-\sigma^2}\,d\sigma\,dx = \left\{ x \int_{\surd x}^\infty e^{-\sigma^2}\,d\sigma \right\}_0^\infty - \int_0^\infty x \frac{d}{dx} \int_{\surd x}^\infty e^{-\sigma^2}\,d\sigma\,dx$$

$$= 0 + \tfrac{1}{2} \int_0^\infty x^{\frac{1}{2}} e^{-x} dx = \tfrac{1}{4}\pi^{\frac{1}{2}},$$

by partial integration.

Consequently

$$\mathrm{A} = 0,$$

that is, the complementary function wholly disappears from χ (and ψ); moreover, the differential equation for χ is satisfied whatever the value of C. But the preceding integrations show also that

$$\frac{1}{\tau}\overline{[\rho^2\chi]} = \frac{1}{\tau_0}\overline{[\rho_0{}^2\chi_0]}, \quad \frac{1}{\tau}\overline{[\rho^2\psi]} = \frac{1}{\tau_0}\overline{[\rho_0{}^2\psi_0]},$$

and therefore (by 29–31)

$$\bar{\xi}/\tau = \mathrm{C}\frac{\partial\alpha}{\partial x} + \mathrm{C}'\,\mathrm{X},$$

or

(36) $$\bar{\boldsymbol{\rho}}/\tau = \mathrm{C}\frac{\partial\alpha}{\partial\mathbf{r}} + \mathrm{C}'\,\mathbf{F}.$$

8. Thus C and C′ determine the mean velocities of the grains due to the non-uniformity of the fluid and to the external forces applied to the grains. These constants may have any values so far as concerns the general differential equation (13) satisfied by ϕ; their values will depend on the physical conditions in each particular case, and must be determined by other considerations than those merely of conservation of grains, on which (13) is based. The value of C′ in the case when the fluid is a liquid is, in fact, known from Stokes' formula to be

(37) $$1/6\pi a\mu,$$

where a is the radius of the grain and μ is the coefficient of viscosity of the liquid; when the fluid is a gas in which the mean free path is not vanishingly small compared with a, this formula for C′ has to be corrected in a way which has been discussed by Cunningham, Millikan and others.

Einstein has shown that the coefficient of diffusion D of the grains is simply related to the constant C′, *i.e.*,

(38) $$\mathrm{D} = k\mathrm{T}\,\mathrm{C}',$$

where T denotes the absolute temperature of the fluid, and k is Boltzmann's gas constant, $1 \cdot 37 . 10^{-16}$. Hence the function ψ, given by the same equation as χ, *i.e.*, (35), except for the omission of the term containing $d \log \mathrm{D}/d\alpha$, is (since $\mathrm{A} = \mathrm{B} = 0$)

$$\psi = \mathrm{C}'/2\mathrm{D} = 1/2k\mathrm{T}. \tag{39}$$

Unless the non-uniformity of the fluid is such that it produces a resultant force on the grains, it would appear that the constant C must be zero. If, as we suppose, the fluid is at rest, the pressure will increase uniformly downwards, owing to gravity; the resultant pressure of the fluid on a grain is just sufficient to support a volume of the fluid equal to that of the grain, but it is convenient to subtract this resultant force from the weight of the grain (if the fluid and grains are not of equal density) and to treat the difference as part of the external force **F**. Apart from this the fluid will apparently exert no resultant force, tending to produce a mean motion of the grains, whatever other properties of the fluid, such as temperature or composition, may be non-uniform. If this is so we may write $\mathrm{C} = 0$, and the complete expression for ϕ becomes

$$\begin{aligned} \phi &= \frac{1}{(4\pi \mathrm{D}\tau)^{3/2}} e^{-\rho^2/4\mathrm{D}\tau} \left\{ 1 + \tfrac{1}{4}\left(\frac{\rho^2}{2\mathrm{D}\tau} - 5\right) \frac{d \log \mathrm{D}}{d\alpha} \left(\boldsymbol{\rho}, \frac{\partial \alpha}{\partial \mathbf{r}}\right) + \frac{1}{2k\mathrm{T}} (\boldsymbol{\rho}, \mathbf{F}) \right\} \\ &= \frac{1}{(4\pi \mathrm{D}\tau)^{3/2}} e^{-\rho^2/4\mathrm{D}\tau} \left\{ 1 + \tfrac{1}{4}\left(\frac{\rho^2}{2\mathrm{D}\tau} - 5\right) \left(\boldsymbol{\rho}, \frac{\partial \log \mathrm{D}}{\partial \mathbf{r}}\right) + \frac{1}{2k\mathrm{T}} (\boldsymbol{\rho}, \mathbf{F}) \right\}. \end{aligned} \tag{40}$$

The expression for ϕ on the latter line is correct however many are the non-uniform properties on which the coefficient of diffusion D depends; and properties which do not affect D do not affect ϕ. The force **F** is supposed to include the gravitational force due to the difference between the densities of the fluid and the grains.

It should be possible to test this distribution law by the method that has been used to verify the symmetrical distribution for a uniform fluid in which the grains are subject to no external force. The natural plan would be to test separately the two additional terms in ϕ, since they are independent. For example, if the grains have the same density as the fluid (which must consequently be a liquid), and no non-gravitational forces are present, $\mathbf{F} = 0$; if also the properties of the fluid depend on one co-ordinate only, say z, and vary uniformly with z, we have

$$\phi = \frac{1}{(4\pi \mathrm{D}\tau)^{3/2}} e^{-\rho^2/4\mathrm{D}\tau} \left\{ 1 + \tfrac{1}{4}\zeta \left(\frac{\rho^2}{2\mathrm{D}\tau} - 5\right) \frac{d \log \mathrm{D}}{dz} \right\}.$$

This implies that fewer displacements fall within the (steadily expanding) circle $\rho^2 = 10\mathrm{D}\tau$ on the side of the plane $z = 0$ on which D is greater, than in the case when D is uniform; outside this circle, and on the same side of the plane

$z = 0$, there are more displacements than in the uniform case. On the other side of $z = 0$, where D has smaller values than at $z = 0$, these relations are reversed. The percentage differences between the non-uniform and uniform cases are represented by the term

$$\tfrac{1}{4}\,\zeta\left(\frac{\rho^2}{2D\tau} - 5\right)\frac{d \log D}{dz},$$

and under possible experimental conditions may be quite appreciable. For example, if a (the radius of a grain) is 10^{-4} cm., and μ, the viscosity of the liquid, is about 0·01, D is of order 10^{-9} at about 300° absolute. Considering the distribution of the displacements during an interval (τ) of 100 seconds, so that the radius of the above circle is 10^{-3} cm., then at $z = \rho = 4 \cdot 10^{-3}$ cm. the percentage difference is

$$\frac{15}{2}\,\frac{d \log D}{dz}.$$

It seems quite possible to make $d \log D/dz$ exceed unity by making the temperature decrease vertically upwards so that within 1 cm. distance D decreases in the ratio $e : 1$; there are many organic liquids for which such a variation in $1/\mu$ (on which D largely depends) occurs with a temperature difference of only about 30° C.

By plotting the points corresponding to the displacements of a large number of grains in a given interval, as in the diagram (suggested by Langevin) given on p. 169 of Perrin's book 'Les Atomes,' the deviation from the normal error-distribution should in suitable cases be quite apparent; and still more so if a second diagram were constructed therefrom, showing the distribution of excess and defect of displacements at different points of the area, compared with the normal distribution for a uniform fluid: the latter diagram should consist of four alternating regions of excess and deficiency, separated, as above described, by the circle $\rho^2 = 10D\tau$ and the line $z = 0$.

Another way of testing the modified distribution function is to determine $\overline{\zeta^3}$; as in the uniform fluid (when no external forces act) $\overline{\xi} = \overline{\eta} = \overline{\zeta} = 0$, and all the mean values of even functions of ξ, η, ζ are the same as in the corresponding uniform case; but whereas in the uniform case the mean value of *every* odd function of ξ, η, ζ vanishes, this is not now true for functions of degree higher than 1, odd only with respect to ζ; in fact,

$$\overline{\zeta^3} = 6\,D^2\tau^2\,\frac{d \log D}{dz},$$

$$\overline{\zeta\xi^2} = \overline{\zeta\eta^2} = \tfrac{5}{4}\,D^2\tau^2\,\frac{d \log D}{dz},$$

and results for higher powers can easily be worked out. With the above values of D, τ, and $d \log D/dz$, *i.e.*, 10^{-9}, 10^2, and 1, we find

$$(\overline{\zeta^3})^{1/3} = 4 \cdot 10^{-5} \text{ cm.}$$

approximately; by making D and τ somewhat larger it should be possible to determine $\overline{\zeta^3}$ experimentally, from the mean of a considerable number of measures of ζ.

The Velocity of Diffusion in the Non-Uniform Case.

9. The velocity with which the grains will move through the fluid under the action of given external forces has already been mentioned, depending on the coefficient C′, which is given approximately by Stokes' formula, or more exactly, in the case of gases, by Millikan's formula. In the present section we consider the velocity of diffusion of the grains due to a non-uniformity of distribution of the grains, or to a non-uniformity of D, or both.

For simplicity suppose that the system is uniform with respect to the coordinates x and y, and varies only with respect to z. Let w be the velocity of diffusion of the grains across a given plane of constant z, so that in an interval τ_0 the net number which cross in the positive direction is $n\tau_0 w$. The value of $n\tau_0 w$ can be calculated by considering where, at the beginning of the interval τ_0, the grains were which in this interval cross the plane. Thus

$$(41) \quad n\tau_0 w = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty n(x, y, z-\zeta, t-\tau_0) \int_\zeta^\infty \phi(\xi, \eta, \zeta', \tau_0, x, y, z-\zeta, t-\tau_0)\, d\zeta'\, d\xi\, d\eta\, d\zeta$$
$$- \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty n(x, y, z+\zeta, t-\tau_0) \int_\zeta^\infty \phi(\xi, \eta, -\zeta', \tau_0, x, y, z+\zeta, t-\tau_0)\, d\zeta'\, d\xi\, d\eta\, d\zeta,$$

where the components of the independent variables $\mathbf{p}$, $\mathbf{r}$ are now shown explicitly. Expanding the complete integrands in powers of ζ and, as in former cases, ignoring terms involving powers higher than the second, because of the powers of τ_0 which they contain after integration, we have

$$(42) \quad n\tau_0 w = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_\zeta^\infty \Big[\Big\{ n\phi_0(\zeta') - \zeta \frac{\partial n\phi_0(\zeta')}{\partial z} + \tfrac{1}{2}\zeta^2 \frac{\partial^2 n\phi_0(\zeta')}{\partial z^2} - \tau_0 \frac{\partial n\phi(\zeta')}{\partial t} \dots \Big\}$$
$$- \Big\{ n\,\phi_0(-\zeta') + \zeta \frac{\partial n\phi_0(-\zeta')}{\partial z} + \tfrac{1}{2}\zeta^2 \frac{\partial^2 n\phi_0(-\zeta')}{\partial z^2} - \tau_0 \frac{\partial n\phi(-\zeta')}{\partial t} + \dots \Big\}\Big]\, d\zeta'\, d\mathbf{p},$$

where the variables in n are x, y, z, t and in ϕ_0 are $x, y, z, \xi, \eta, \zeta, t$. Integrating by parts

$$\int_0^\infty \int_\zeta^\infty n\, \phi_0(\zeta')\, d\zeta'\, d\zeta = n \left\{ \left[\zeta \int_\zeta^\infty \phi_0(\zeta')\, d\zeta' \right]_0^\infty + \int_0^\infty \zeta\, \phi_0(\zeta)\, d\zeta \right\}$$

$$= n \int_0^\infty \zeta\, \phi_0(\zeta)\, d\zeta.$$

Similarly

$$\int_0^\infty \int_\zeta^\infty n\, \phi_0(-\zeta')\, d\zeta'\, d\zeta = -n \int_{-\infty}^0 \zeta\, \phi_0(\zeta)\, d\zeta.$$

Transforming the other terms in like manner, we get

$$(42) \qquad n \tau_0 w = n \bar{\zeta} - \tfrac{1}{2} \frac{\partial}{\partial z} (n \overline{\zeta_0^2}),$$

where n and $\overline{\zeta_0^2}$ refer to $\mathbf{r}$, t. Since we have seen that, if $C = 0$, and in the absence of external forces, $\bar{\zeta} = 0$, this reduces to

$$(43) \qquad w = -\frac{1}{n} \frac{\partial}{\partial z} \left(\frac{\overline{\zeta_0^2}}{2\tau_0} n \right) = -\frac{1}{n} \frac{\partial}{\partial z} (n\mathrm{D}).$$

This is true whether the fluid is uniform or not; from it the equation of continuity (7, § 3) can be at once derived. When the fluid is uniform, so that D is constant, (43) becomes the ordinary equation of diffusion depending on a gradient of concentration of the grains, viz.,

$$w = -\frac{\mathrm{D}}{n} \frac{\partial n}{\partial z};$$

but the appearance of D within the differential coefficient in (43) indicates that even when n is uniform there will be diffusion of the grains if the fluid is not uniform, *i.e.*,

$$w = -\frac{\mathrm{D}}{n} \frac{\partial n}{\partial z} - \frac{\partial \mathrm{D}}{\partial z}.$$

The general expression for the velocity of diffusion $\mathbf{v}$, when the fluid is non-uniform in any way, and the grains are subject to external forces $\mathbf{F}$, is similarly

$$(44) \qquad \mathbf{v} = \frac{\overline{\boldsymbol{\rho}_0}}{\tau_0} + \frac{\mathrm{D}}{k\mathrm{T}} \mathbf{F} - \frac{1}{n} \frac{\partial (n\mathrm{D})}{\partial \mathbf{r}}$$

$$= \frac{\overline{\boldsymbol{\rho}_0}}{\tau_0} - \frac{\mathrm{D}}{k\mathrm{T}} \frac{\partial \mathrm{V}}{\partial \mathbf{r}} - \frac{1}{n} \frac{\partial (n\mathrm{D})}{\partial \mathbf{r}},$$

if the external forces have a potential V.

10. In a steady state of no diffusion $\mathrm{v} = 0$, and if also $\overline{\boldsymbol{\rho}_0} = 0$,

$$\frac{1}{kT}\frac{\partial V}{\partial \mathbf{r}} - \frac{1}{nD}\frac{\partial (nD)}{\partial \mathbf{r}} = 0\,; \tag{45}$$

if the fluid is uniform, so that D and T are constant, this has the well-known solution

$$n \propto e^{-V/kT}. \tag{46}$$

If, however, $V = 0$, as is the case when the grains are of the same density as the fluid, and there are no non-gravitational forces,

$$n\overline{\boldsymbol{\rho}_0} - \frac{\partial (nD)}{\partial \mathbf{r}} = 0,$$

or, if $C = 0$ (§ 8), so that $\overline{\boldsymbol{\rho}_0} = 0$,

$$n \propto 1/D. \tag{47}$$

The variation of D may be due to non-uniform composition of the fluid, but the most important case is that in which it is due to non-uniform temperature. If Stokes' formula is applicable,

$$n \propto 6\pi a\mu/kT,$$

or, ignoring the small variation in the dimensions of the grains,

$$n \propto \mu/T.$$

For many gases $\mu \propto T^s$, where $\frac{1}{2} < s < 1$, so that in this case $n \propto T^{s-1}$; thus the grains would in the steady state be more concentrated in the colder regions of the gas. The case $s = 1$ corresponds to a Maxwellian gas, in which the molecules repel one another with forces varying inversely as the fifth inverse power of the mutual distance; in such a gas $n \propto T^{1-1}$, *i.e.*, n is constant; it is curious that in such a gas "Brownian" thermal diffusion ceases to operate just as does true gaseous thermal diffusion.

In the case of liquids μ decreases with increasing temperature, in some cases very rapidly; for example, in glycerine $\mu = 46$ at 0° C., and at 30° C. it drops to 3·5. Thus if a set of Brownian particles of the same density as glycerine were suspended in this fluid, and if a temperature gradient were maintained in the glycerine, T varying from 0° to 30° C., the grains would attain a steady distribution in which they would be more than 10 times as numerous in the cold layer as in the hot. Even in liquids less remarkable than glycerine, the inverse ratio of the concentrations of the grains in the hot (T_1) and cold (T_2) regions always exceeds T_1/T_2, whereas in gases it is always less than (T_1/T_2). Thus Brownian thermal diffusion is more powerful in liquids than in gases.

11. For a fluid in which there is a uniform vertical temperature gradient, it is possible to determine the steady distribution of the grains even when their density is not the same as that of the fluid. If the mass of a grain, minus that of the fluid which it displaces, is m' (regarded as sensibly constant), the external force (reckoned positive upwards) is $-m'g$. Thus, putting $v = 0$ in (44), the equation of equilibrium is

$$\frac{\mathrm{D}m'g}{k\mathrm{T}} + \frac{1}{n}\frac{\partial(n\mathrm{D})}{\partial z} = 0,$$

and we have also

$$\frac{d\mathrm{T}}{dz} = \gamma,$$

where γ is the temperature gradient. Hence

$$\frac{\gamma}{n\mathrm{D}}\frac{d(n\mathrm{D})}{d\mathrm{T}} = -\frac{m'g}{k\mathrm{T}} = \quad \frac{m'g}{k}\frac{d\log\mathrm{T}}{d\mathrm{T}}.$$

Hence

$$(n\mathrm{D})^{\gamma}\,\mathrm{T}^{m'g/k} = \text{constant},$$

or

$$n \propto \mathrm{D}^{-1}\,\mathrm{T}^{-m'g/k\gamma},$$

whereas by (47), when there is no temperature gradient,

$$n \propto e^{-m'gz/k\mathrm{T}}.$$

If the grains are heavier than the fluid, and T increases upwards, so that m' and γ are both positive, n increases downwards still more rapidly than when (47) applies. On the other hand, by suitably choosing the temperature gradient and the density of the grains (in this case *less* than that of the fluid), so that

$$\mathrm{D} \propto \mathrm{T}^{-m'g/k\gamma},$$

a uniform concentration may be maintained by a balance between the opposing effects of thermal diffusion and gravitational action.

12. It is worth while mentioning explicitly that so long as we neglect the small difference in size of the grains between the hot and cold regions, due to their thermal expansion, and so long as Stokes' law is valid, the concentration ratio of the grains as between the hot and cold regions, due to Brownian thermal diffusion alone (uncomplicated by gravitational settling), is *independent of the size of the grains* so long as $\mathrm{C} = 0$ (§ 8): for $n_1/n_2 = \mathrm{D}_2/\mathrm{D}_1 = (\mathrm{T}_2/\mathrm{T}_1)(\mu_1/\mu_2)$. It would be possible to show this experimentally if a horizontal temperature gradient could be maintained without convection currents: the horizontal concentration-gradient due to pure Brownian thermal diffusion would then be

independent of the vertical concentration gradient due to gravitational settling.

The time required for the attainment of the steady state will, of course, depend on D, and therefore on the size of the grains, being inversely proportional to D and therefore directly proportional to a.

Summary.

(1) The distribution in magnitude and direction of the Brownian displacements, in any assigned interval, of small spherical grains suspended in a fluid, liquid or gaseous, is obtained for the general case when the fluid is non-uniform in temperature, composition, or any other property which affects the coefficient of diffusion of the grains (§ 8). It includes an undetermined constant, not zero in general, but which is assumed to be zero for Brownian grains large compared with the mean free path of the molecules of the fluid.

(2) The steady distribution of grains in a fluid is determined when the temperature T of the fluid is non-uniform, both when the grains are of the same density as the fluid, and when there is a difference of density, so that the grains are acted on by differential gravity. In the former case the concentration n is found to vary inversely as the coefficient of diffusion D; D varies with the temperature more rapidly in liquids than in gases, and in the case of a Maxwellian gas D is independent of temperature, so that "Brownian" thermal diffusion, like true gaseous thermal diffusion, does not operate in such a gas. In other cases the grains are most concentrated in the coldest regions. When the grains have not the same density as the fluid, it is shown that with a uniform temperature gradient nD is not constant, as in the former case, but varies as a power of T.

(3) The order of magnitude of the various effects is estimated numerically, and it is shown that they should be susceptible of experimental test.

The Molecular Displacements in Diffusing Gas-Mixtures.

By S. CHAPMAN, M.A., D.Sc., F.R.S.

(Received March 7, 1928.)

(1) In the preceding paper the principal arguments are of a general nature, not confined to the case of Brownian particles suspended in a medium; they can be applied to ordinary gas-mixtures. In particular, the equations (7), (13), (44) do not depend on the nature and relative number of the particles considered, nor on the nature of the surrounding medium, provided that (i) the particles retain their individuality, neither breaking up nor combining with others, (ii) no boundary effects are considered, and (iii) a small minimum interval τ_0 exists such that in successive intervals τ_0 the displacements of any particle are independent. In gases the frequency of molecular collisions is so great that the existence of such an interval is almost self-evident. Its value depends on the extent to which molecular velocities persist after collision.* For the molecules in air, under normal conditions, τ_0 will not exceed 10^{-7} second.

It is of interest, both from the standpoint of the theory of gases, and as throwing light on the preceding paper, to deal briefly with the application of the above equations to gas-mixtures. The equations of the present note will be numbered in succession to those of the preceding paper.

(2) In a uniform gas-mixture Einstein's displacement-distribution function (23) is valid for each kind of molecule, whatever their relative numbers. There will be a function ϕ and a constant D for each kind of molecule. These constants measure the rate at which the individual molecules of either kind diffuse through the mixture; they may be called coefficients of partial diffusion (D_1, D_2) to distinguish them from the ordinary coefficient of (mutual) diffusion (D_{12}); in general D_1, D_2 and D_{12} are unequal.

When the molecules of one kind (say (1)) are very few in number compared with those of kind (2), and when the "medium" formed by the latter molecules is uniform, the coefficient D_1 can be identified with D_{12}; for in this case each molecule of kind (1) will be independent of its distant neighbours of its own kind, and will be surrounded by the uniform medium (2), so that D_1 will be independent of position, and $\bar{\rho}_1$ will by symmetry be zero. Hence (44) becomes (taking F or V to be zero)

$$v_1 = -\frac{1}{n_1}\frac{\partial(n_1 D_1)}{\partial r} = -\frac{D_1}{n_1}\frac{\partial n_1}{\partial r},$$

---

* *Cf.* Jeans, 'Dynamical Theory of Gases,' pp. 260–267, 4th edn.

which is the equation of diffusion relative to the stationary medium. Hence in this case $D_1 = D_{12}$.

Likewise in this case (n_1/n_2 small) D_2 can be identified with the coefficient of self-diffusion D_{22} of the pure gas (2). Thus we have, when

$$n_1/n_2 \to 0, \qquad D_1 = \lim_{n_1/n_2 \to 0} D_{12}, \qquad D_2 = D_{22},$$

and similarly when

$$n_2/n_1 \to 0, \qquad D_1 = D_{11}, \qquad D_2 = \lim_{n_2/n_1 \to 0} D_{12}.$$

This shows that D_1 and D_2 are in general unequal to each other and to D_{12}, and that they depend on the concentration-ratio n_1/n_2.

The coefficient D_{12} depends essentially on the interaction between the two kinds of molecules, while D_{11} and D_{22} each depend on the properties of only one kind of molecule. For a perfect gas composed of Maxwellian molecules, obeying the fifth-power law, D_{12} is independent of the ratio n_1/n_2, but in other cases it is variable with n_1/n_2, though only by a few per cent. Ignoring this variation, and regarding the molecules as rigid elastic spheres of masses m_1, m_2 and diameter d_1, d_2, then if the mixture is a gas at an assigned pressure and temperature,

$$D_{11} : D_{12} : D_{22} :: \frac{1}{m_1^{\frac{1}{2}} d_1^2} : \left\{2\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\right\}^{\frac{1}{2}} \left(\frac{2}{d_1 + d_2}\right)^2 : \frac{1}{m_2^{\frac{1}{2}} d_2^2}.$$

Thus as the proportion of the molecules (¹) in the mixture varies from 0 to 1, D_1 varies between initial and final values which are in the ratio

$$\left(\frac{d_1 + d_2}{2d_1}\right)^2 \left\{\frac{2m_2}{m_1 + m_2}\right\}^{\frac{1}{2}}.$$

So far as I know, no theory yet exists which indicates the variations of D_1 and D_2, between their limiting values for n_1/n_2 zero and infinite.

(3) It is of interest to point out the distinction between the velocities $\mathbf{v}_1$ and $\boldsymbol{\rho}_1/\tau$ at any point instant $\mathbf{r}$, t; the component of the former in any direction is the mean velocity with which the particles (¹) cross a plane through $\mathbf{r}$ normal to that direction; the latter is the mean velocity of the cluster of molecules (¹) in a small volume element around $\mathbf{r}$ at time t. In general these two velocities are unequal, as appears from (44).

(4) The general form of the displacement-distribution function ϕ obtained in § 7 of the preceding paper for the case where there is no external force, but the medium is non-uniform with respect to certain scalar properties α, is

$$(48) \quad \phi(\boldsymbol{\rho}, \tau) = \frac{1}{(4\pi D\tau)^{3/2}} e^{-\rho^2/4D\tau}\left[1 + \sum_\alpha \left\{\frac{C_\alpha}{2D} + \frac{1}{4D}\frac{\partial D}{\partial \alpha}\left(\frac{\rho^2}{2D\tau} - 5\right)\right\}\left(\boldsymbol{\rho}, \frac{\partial \alpha}{\partial \mathbf{r}}\right)\right],$$

where the C_α's are constants whose values cannot be determined merely by considering the conservation of grains, without reference to the physical circumstances. As in (30), it is easily deduced that

$$(49) \qquad \frac{\bar{\boldsymbol{\rho}}}{\tau} = \sum_\alpha C_\alpha \frac{\partial \alpha}{\partial \mathbf{r}}.$$

One important scalar property of a mixture is the ratio of the two kinds of particles; it is convenient to write

$$(50) \qquad n_0 = n_1 + n_2, \qquad n_{10} = n_1/n_0, \qquad n_{20} = n_2/n_0.$$

so that

$$(51) \qquad n_{10} + n_{20} = 1.$$

Suppose the non-uniform property α is taken to be n_{10}, and let the coefficients C_α in ϕ_1 and ϕ_2 be denoted in this case by A_1 and A_2, so that

$$(52) \qquad \frac{\bar{\boldsymbol{\rho}}_1}{\tau} = A_1 \frac{\partial n_{10}}{\partial \mathbf{r}}, \qquad \frac{\bar{\boldsymbol{\rho}}_2}{\tau} = A_2 \frac{\partial n_{10}}{\partial \mathbf{r}}.$$

Then we have

$$(53) \qquad \mathbf{v}_1 = A_1 \frac{\partial n_{10}}{\partial \mathbf{r}} - \frac{1}{n_1} \frac{\partial (n_1 D_1)}{\partial \mathbf{r}}, \qquad \mathbf{v}_2 = A_2 \frac{\partial n_{10}}{\partial \mathbf{r}} - \frac{1}{n_2} \frac{\partial (n_2 D_2)}{\partial \mathbf{r}}.$$

If the gas is at rest as a whole,

$$(54) \qquad n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2 = 0,$$

whence we obtain

$$(55) \qquad A_1 n_1 + A_2 n_2 = \frac{d(n_1 D_1 + n_2 D_2)}{dn_{10}}.$$

Moreover, comparing (53) with the equation of diffusion for either kind of particle, which, when (53) is satisfied, is

$$(56) \qquad \mathbf{v}_1 = -\frac{D_{12}}{n_{10}} \frac{\partial n_{10}}{\partial \mathbf{r}}, \qquad \mathbf{v}_2 = -\frac{D_{12}}{n_{20}} \frac{\partial n_{20}}{\partial \mathbf{r}},$$

we have

$$(57) \qquad D_{12} = \frac{1}{n_0} \frac{d(n_1 D_1)}{dn_{10}} - n_{10} A_1 = -\frac{1}{n_0} \frac{d(n_2 D_2)}{dn_{10}} + n_{20} A_2$$
$$= \frac{1}{2n_0} \left\{ \frac{d(n_1 D_1 - n_2 D_2)}{dn_{10}} - (n_{10} A_1 - n_{20} A_2) \right\}.$$

The equations (55), (56) enable A_1 and A_2 to be determined if D_1, D_2, D_{12} are known, or, conversely, if A_1 and A_2, as well as D_{12}, could be calculated, D_1 and D_2 would be determinate. At present, however, no method of calculating A_1, A_2, D_1, D_2 appears to be known, except when n_1/n_2 is very small or very large.

(5) We now return to the case of a perfect gas-mixture in which the concentration of the molecules $(^1)$ is very small; the independent variables of state of the mixture may be taken to be n_{10}, the pressure p, and the temperature T.

Let us suppose that p is uniform but that T is non-uniform; it is immaterial whether or not n_{10} is uniform so long as it is small, because in this case A_1 (§ 4) is zero. It is convenient in this case to identify α with log T, and to write C_T for the corresponding value of C_α. Then, by (48), identifying D with D_{12} as before, we have

$$\mathbf{v}_1 = C_T \frac{\partial \log T}{\partial \mathbf{r}} - \frac{1}{n_1} \frac{\partial (n_1 D_{12})}{\partial \mathbf{r}}. \tag{58}$$

But from the known theory of diffusion in gases we have, relative to the mean velocity of the medium,

$$\mathbf{v}_1 = -\frac{D_{12}}{n_{10}} \left\{ \frac{\partial n_{10}}{\partial \mathbf{r}} + k_T \frac{\partial \log T}{\partial \mathbf{r}} \right\}, \tag{59}$$

where $k_T D_{12}$ is the coefficient of thermal diffusion (D_T). If the pressure is constant, $n_0 \propto 1/T$, so that (59) may be written

$$\mathbf{v}_1 = -\frac{D_{12}}{n_1} \frac{\partial n_1}{\partial \mathbf{r}} - D_{12} \left\{ 1 + \frac{k_T}{n_{10}} \right\} \frac{\partial \log T}{\partial \mathbf{r}}. \tag{60}$$

Equating the two values of $\mathbf{v}_1$ given by (58) and (60), and cancelling out common terms and factors, we have

$$\frac{C_T}{D_{12}} = \frac{d \log D_{12}}{d \log T} - \left(1 + \frac{k_T}{n_{10}} \right). \tag{61}$$

Thus, while we have no direct method of determining C_T in general, the theory of gaseous diffusion enables C_T to be deduced for a gas-mixture of small concentration.

In the case of Maxwellian molecules $k_T = 0$ and $D_{12} \propto T^2$, so that

$$C_T / D_{12} = 1.$$

This implies that the molecules of kind $(^1)$ which start from any given region at time t have a mean velocity, in the direction of increasing temperature, of amount

$$\frac{D_{12}}{T} \frac{\partial T}{\partial r},$$

since $\bar{\boldsymbol{\rho}}/\tau = C_T \, \partial \log T / \partial \mathbf{r}$; this result is true whether the molecules $(^1)$ are lighter or heavier than the molecules $(^2)$. The displacement-distribution function in this case is

$$\phi(\boldsymbol{\rho}, \tau) = \frac{1}{(4\pi D\tau)^{3/2}} e^{-\rho^2/4D\tau} \left\{ 1 + \left(\frac{\rho^2}{4D\tau} - \frac{3}{2} \right) \left(\boldsymbol{\rho}, \frac{\partial \log T}{\partial \mathbf{r}} \right) \right\}.$$

In the case of molecules which are rigid elastic spheres, $D_{12} \propto T^{3/2}$, so that

$$\frac{C_T}{D_{12}} = \frac{1}{2} - \frac{k_T}{n_{10}}.$$

The expression for k_T, even when n_1/n_2 is very small, is a complicated function of the molecular masses and diameters, but one or two special cases may be noted. If m_2/m_1 is large, that is, if the rare molecules are much lighter than the abundant ones, $k_T/n_{10} = -\frac{1}{2}$; hence again we have

$$C_T = D_{12}.$$

If, however, m_1/m_2 is large, the value of k_T/n_{10} is (to a first approximation)

$$\frac{5}{4\sqrt{2}}\left(\frac{d_1 + d_2}{2d_2}\right)^2,$$

so that

$$\frac{C_T}{D_{12}} = \frac{1}{2} - \frac{5}{4\sqrt{2}}\left(\frac{d_1 + d_2}{2d_2}\right)^2.$$

This is zero if $d_1/d_2 = \sqrt{(\frac{8}{5}\sqrt{2})} - 1$, or $\frac{1}{2}$ (very nearly); for larger values of d_1/d_2 it is negative, and may become large; for very small values of d_1/d_2 its lower limit (when $d_1 = 0$) is $0 \cdot 23$.

Thus for a gas-mixture in which the concentration of molecules (1) is small, C_T may be positive, negative, or zero. If the molecules are rigid elastic spheres, and if the rare molecules are much the lighter, any particular cluster of molecules will, while spreading outwards, move steadily towards the hot region; but if the rare molecules are much the heavier they will, unless their diameter is small (less than half that of the molecules (2)), spread outwards and at the same time travel towards the colder regions. The spreading out is not symmetrical relative to the mean motion, on account of the last term in the function ϕ (*cf.* (48)).

(6) In conclusion, we consider the steady state in the presence of a temperature gradient. Writing $v_1 = 0$ in (58) and (60), we have

$$n_1 \propto T^{C_T/D_{12}}/D_{12}, \qquad n_1 \propto T^{-(1+k_T/n_{10})} \tag{62}$$

which are in agreement, by (61). In the preceding paper I suggested that for grains large compared with the mean distance between the molecules (2) $C_T = 0$, so that $n_1 \propto 1/D_{12}$; though the conditions in this case are unlike those considered in the theory of gaseous diffusion, it is of interest to see how far the conclusion $n_1 \propto 1/D_{12}$ is departed from in the gaseous case. For Maxwellian molecules, whatever the masses or force-constants, $C_T/D_{12} = 1$, so that the conclusion $n_1 \propto 1/D_{12}$ (or $n_1 \propto T^{-2}$) is wrong by a factor T, *i.e.*, actually in this case $n_1 \propto T^{-1}$. For rigid elastic spheres, when m_1/m_2 is large and $d_1 > \frac{1}{2}d_2$, C_T/D_{12} is negative, and therefore the concentration-gradient set up by thermal diffusion *exceeds* that corresponding to $n \propto 1/D_{12}$.

Thus in gas-mixtures of small concentration there may be deviations from the result $n \propto D_{12}^{-1}$ in either direction. At present it is impossible to give a definite theoretical prediction as to how large grains in a gas (and still less in a liquid) will behave in respect to thermal diffusion.

The Excitation of Soft X-Rays.—II.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society, and F. C. Chalklin, Ph.D.

(Received March 12, 1928.)

1. *Additional Observations on Iron.*

This paper is a continuation of our former paper[*] with the same title. In our work on iron we were compelled, for technical reasons, into which it is unnecessary to enter, to leave a gap between about 130 and 200 volts. We have since completed that investigation by filling up the gap and have also extended the range from 600 to 720 volts. The first section of the present paper describes the results obtained.

The steady deflection method was used and only two important changes were made in the apparatus. By using a new suspension of Wollaston wire for the electrometer, it was found possible to increase the sensibility to 2400 mm. per volt with little loss of steadiness. This enabled either lower filament currents to be used with less consequent likelihood of gas troubles or alternatively to work with a leak of lower resistance across the electrometer quadrants.

A new high resistance was therefore constructed having a smaller resistance than the original xylol-alcohol tube. This was made by depositing lamp black on a quartz rod, the contacts on to wires wound round the ends of the rod being made with india ink.

[*] 'Roy. Soc. Proc.,' A, vol. 110, p. 247 (1926).

The rod was then sealed up in a brass tube provided with an attachment containing phosphorous pentoxide. The lead from one end of the rod was insulated from the brass by a quartz tube. This type of resistance has the merit of being practically non-polarising and does not appear to possess the large temperature coefficient of the xylol-alcohol mixtures.

We observed three discontinuities between 600 and 720 volts. Owing to the large amount of general radiation at these high voltages, the discontinuities are not well marked and are only to be observed under very good experimental conditions. For the same reason their exact position is difficult to locate. Our final corrected mean values are given in Table I together with all the values recorded by other observers in this region.

Table I.

Observer.	Critical Voltage.		
R. and C.	608·1	640	689·7
Thomas	—	639	704·3
Rollefson	618	637	697

The point at 608·1 was observed on four curves out of six, the one at 640 on one curve out of six and the one at 689·7 on four out of nine. This statement probably gives a fair idea of the intensity of the effects. No particular weight should be attached to the figure 640, which rests on a single observation. The figure 608·1 may be about 3 volts too low. All these figures lie just under the L_{III} level for iron, which is about 706 volts.

We found eight discontinuities between 130 and 200 volts. A number of them are shown in fig. 1. As these are so very close together a great many observations had to be taken to locate them with accuracy. The results of all the experiments are given in Table II. It was not possible to cover the whole of this range in one experiment, and in each horizontal row the experiment included only the range indicated by the figures on the extreme right and left respectively. Where there is a gap within this range it means that the point in question was either absent or not sufficiently clearly marked to base an estimation on. After the first 10 experiments the filament was replaced by a shorter one with a view to improving the resolution. Experiments 17, 18, and 19 were unsatisfactory and have not been used.

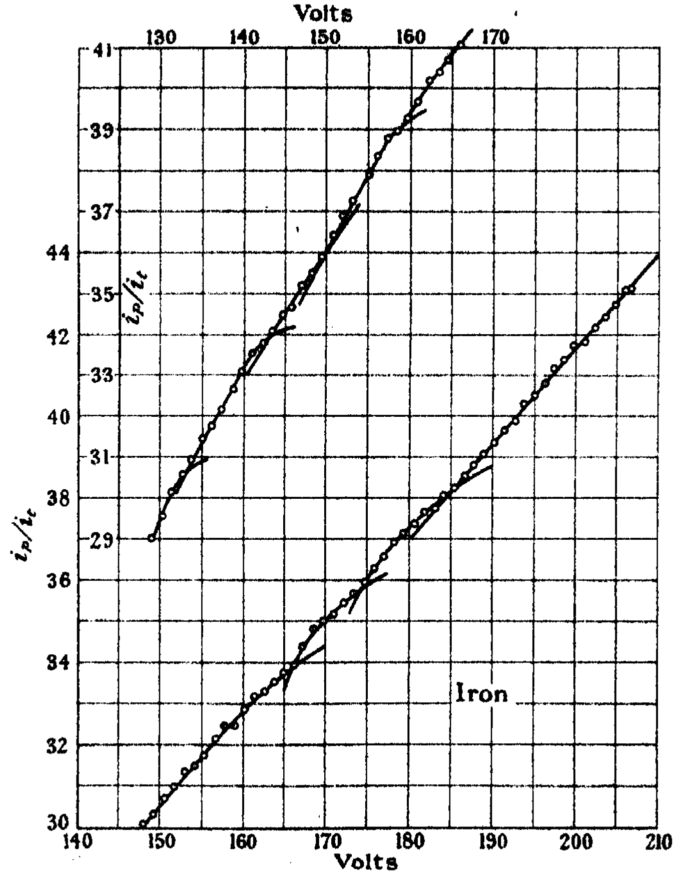


Fig. 1.

Table II.—Critical Voltages for Iron between 130 and 200 Volts.

No. 1	10.11.25		191·5		176·1				
,, 2	19.11.25	194·7							
,, 3	12.1.26		188·9		177·1	168·6			
,, 4 ...	12.1.26	195·7	190·7			163·2			
,, 5	20.1.26		190·1		178·1	doubtful	155·3	146·5	
,, 6	20.1.26		189·2		175·6	,,	157·1		
,, 7 ...	21.1.26			183·3	173·5	165·7	155·5		
,, 8	25.1.26	198·6	192·1		175·6	164·6			
,, 9	25.1.26		188·6			167·6	155·2		140·0
,, 10	25.1.26					169·7	154·7		
,, 11	26.2.26	196·1	186·6		174·6				
,, 12 ...	—.3.26	199·4	190·4	182	171·8	165·8			
,, 13	9.3.26				172·8	162·8	155·9		
,, 14	9.3.26	196·8	187·8		175·3	166·3			
,, 15	10.3.26	196	189·8	doubtful	172·3		155·4		
,, 16	11.3.26					162·2	153·7	147·5	136·6
,, 20	17.3.26						158·8		136·7
,, 21 ...	—.5.26						157·8	148·4	137·1
,, 22	—.5.26						158·0	151·5	142·3
Mean		196·8	189·6	(182·7)	174·8	165·6	156·1	148·5	188·5

Table III.

R. and C.	Thomas.	Rollefson.	Kurth.	Chu.	A., D. and H.	Petry.
45·6 (3)	41·2 (2·5 I)					
	48·0 (2·4 R)					43·6 (2)
50·6 (3)	51·3 (1·6 R)	51·7	50		52	46·5 (4)
	54·6 (3·1 I)			57		
63·7 (4)	62·0 (2·7 R)					
	67·6 (1·6 R)					63·8 (3)
75·6 (2)	75·3 (2·2 R)			73		70·4 (3)
					78	
84·8 (2)	82·7 (2·3 R)	86·6	81·7			
	87·? (1·1 R)					
	91·2? (1·3 R)					85·1 (1)
95·0 (3)	94·8 (1·3 R)	100·3	95·4		95	
107·6 (0)	103·5 (2·0 I)					97·6 (4)
115·1 (1)	112·2 (1·7 R)	116·1			117	
	125·7 (1·5 R)					119·4 (1)
131 (1)	131·8 (1·9 R)					
138·5 (3)	136·0 (1·3 R)	135·3				
	140·7 (1·6 R)					
148·5 (1)	150·2 (1·6 R)	152·0				
156·1 (3)	159·8 (1·3 R)	158·3				157·2 (2)
165·6 (2)	169·4 (2·3 I)	165·0			171	
174·8 (3)						
182·0 (1)*	181·6 (1·2 R)					175·6 (2)
189·6 (3)	192·0 (1·8 R)				186	
196·8 (3)						193·3 (?)
	212·0 (1·5 R)					
221·5 (2)	221·3 (1·1 R)				219	
	228·1 (2·8 I)					227·3 (?)
240·6 (3/1)	241·? (1·0 R)					
	250·8 (1·4 R)					
	261·5 (1·1 R)					257·0 (2)
270·2 (2)	271·5 (1·2 R)					
	277·4 (1·4 R)					
288·3 (10)	288·6 (1·1 R)					290·0 (1)
	303·7 (1·8 R)					
	331·0 (1·6 R)					
						348·0 (5)
						390·0 (1)
						455·0 (2)
498·0 (2)						
608·1 (1)		623				
640 (0)†	639·0 (0·9 R)	642				
689·7 (1)	704·3 (1·5 I)	697				720 (1)
	818·5 (2·0 I)					
						1040 (?)
						1370 (?)

* Only observed twice.
† Observed only once.

The following critical potentials have been observed at potentials below the range of the voltage in the table. For soft X-ray excitation Thomas gives 7·3 (2·5 R), 11·1 (2·7 I), 14·1 (2·2 R), 16·5 (2·9 R), 19·4 (1·8 I), 24·3 (1·3 R), 28·8 (2·5 R), 34·3 (2·0 R), 37·5 (1·1 R). For secondary electron emission Petry gives 7·3 (10), 11·6 (10), 14·4 (8), 18·3 (4), 22·6 (4), 25·0 (2), 29·0 (1).

The following are additional critical potentials for iron which have been given by Stuhlman :—3·3, 8·5, 10·4, 24·3, 50·0, 80·0, 120, 160, 171, 200 volts. The values have been taken from the following papers :—

C. H. Thomas, ' Phys. Rev.,' vol. 26, p. 739 (1925) ; G. K. Rollefson, ' Phys. Rev.,' vol. 23, p. 35 (1924) ; E. H. Kurth, ' Phys. Rev.,' vol. 18, p. 461 (1921) ; Chu, ' Journal of the Franklin Institute,' p. 615 (1925) ; Andrewes, Davies and Horton, ' Roy. Soc. Proc.,' A, vol. 117, p. 649, (1928) ; R. L. Petry, ' Phys. Rev.,' vol. 26, p. 346 (1925) ; O. Stuhlman, ' Science,' vol. 56, p. 344 (1922), ' Phys. Rev.,' vol. 25, p. 234 (1925).

The brackets in the table are to indicate that the discontinuity is due to two (in one case three) others unresolved.

In Table III are assembled all the known observations of critical potentials of iron both for soft X-ray excitation got by various observers and also those given by Petry for secondary electron emission. The agreement between different observers for the soft X-ray inflections seems to us very satisfactory when the difficulty of resolving some of these close inflections and the probable presence of some impurity effects are taken into consideration. We have taken the liberty of adding 4·9 volts to all Rollefson's values in order to allow for the work function at the iron target and the contact potential between this electrode and the hot cathode. If the special emission of secondary electrons in Petry's experiments sets in at the same voltages as the special excitation of soft X-rays, his critical voltages should be 4·9 volts lower than the corresponding critical voltages for soft X-ray excitation, this being the magnitude of the work function correction, which there is no obvious reason otherwise for applying to the secondary electron emission data. It will be seen that nearly all Petry's critical potentials are about 5 volts below some soft X-ray critical potential, the only clear exception being the rather strong group 348 (5), 390 (1), 455 (2), to which there is nothing corresponding in the soft X-ray list. It seems probable that the critical soft X-ray potentials, even in the long list given by Thomas, are nearly all real and due to iron. Those marked ? are open to suspicion, but it is doubtful if many of the others are. We are inclined to attribute 331 to an impurity (see para. 3 below).

2. *Series Relations among the Critical Voltages.*

The lower voltage region examined in the present work is important because in it Rollefson claims to have found a number of critical voltages for iron which follow a Rydberg series, that is to say, the voltages can be written in the form

$$V(n) = A - b/n^2 \qquad n = 5, 6, 7, 8, 9, 10, -, 12. \qquad (1.1)$$

This would be an important discovery in any event, but in the present case it is the more remarkable because the value of b given by these values of n corresponds to an attracting centre with about 13 times the electronic charge. There are, of course, enough electrons (26) in the iron atom for this to be possible, but it corresponds to a state of affairs which has not been contemplated in any theory of atomic structure hitherto put forward. In view of the importance of this question we have reconsidered the existing data critically.

In the first place it seemed likely that the resolution obtained both by Thomas and by ourselves was superior to that of Rollefson, so that, in the hope of getting more accurate data, we took the mean of the values got by Rollefson, by Thomas and by ourselves for the inflections which appeared to be the same as Rollefson's series inflections. These means are given in row 2 of Table IV. We then used the numbers for $n = 5$ and $n = 12$ to determine A and b in equation (1.1) and so determined the remaining critical values. These are set out in row 3. It will be seen that the agreement with the mean values in row 2 is very poor. In particular the difference between 136·6 and 131·3 seems beyond the limits of any possible error, and the discrepancy for $n = 8$ is probably outside the limits of error. However, there is a critical potential at 131 (131 (1) R. and C., 131·8 (1·9 R) Thomas).

On further consideration it seemed to us likely that a more accurate basis of comparison would be got by keeping entirely to the data of Thomas, which have all been got by the same technique and with very good resolution, rather than by taking the means of values got by different experimenters by different methods, especially as the identification of some of the corresponding critical voltages is uncertain. We now took Thomas's values of 28·8 and 150·2 as $n = 4$ and $n = 9$ respectively, from which $A = 180 \cdot 0$, $b = 2420$. These constants give the values in row 6. Thomas's nearest critical voltages are given in row 4 and in row 5 his estimates of intensity and character. The number such as 2·5 measures the increase in slope at the critical voltage and the letter R indicates that the steepness diminishes with increasing voltage after passing the critical point: the letter I indicates that it does not. It will be seen that the agreement is now good. At the left-hand side of the table Thomas's critical voltages are about 5 volts apart and on the right-hand side about 10 volts apart. There are no missing breaks until $n = 10$. Furthermore,

Table IV.

1. Fe n =	4	5	6	7	8	9	10	11	12
2. Means		84·2*	116·1	136·6	147·2	151·1	158·1		165·3*
3. Calc.	29·0	84·2*	114·2	131·3	144·0	152·1	157·8	162·1	165·3*
4. Thomas	28·8*	82·7	112·2	131·8	140·7	150·2*		159·8	
5. Int.	(2·5 R)	(2·3 R)	(1·7 R)	(1·9 R)	(1·6 R)	(1·6 R)		(1·3 R)	
6. Calc.	28·8*	83·2	112·7	130·6	142·2	150·1*	155·8	160·0	163·2
7. Thomas		54·6*	82·7	103·5	112·2	125·7		131·8	136·0
8. Int.		(3·1 I)	(2·3 R)	(2·0 I)	(1·7 R)	(1·5 R)		(1·9 R)	(1·3 R)
9. Combination	0·6	54·5*	84·5	103·6	112·7	122·0	(127·6)	131·6	(135·0)
10. Co. Thomas	32·2	84·3	114·3	131·4*	143·8		154·0		163·5
11. Int.	(2·2 R)	(1·9 R)	(1·9 R)	(1·5 R)	(1·5 R)		(1·3 R)		(1·3 R)
12. Calc.	32·2*	85·2	114·0	131·4*	142·7	150·4	155·9	160·1	163·2
13. Thomas		55·8*	84·3	104·4	114·3		125·5	131·4	
14. Int.		(2·9 I)	(1·9 R)	(2·0 I)	(1·9 R)		(1·3 R)	(1·5 R)	
15. Combination	3·7	55·8*	85·5	102·9	115·3	(121·9)	125·5	(131·6)	135·0
16. Ni. Thomas	34·6*	87·2	116·6	113·7*	145·0	156·8			
17. Int.	(2·2 R)	(1·6 R)	(1·2 R)	(1·4 R)	(1·3 R)	(1·4 R)			
18. Calc.	34·6*	87·6	116·4	133·8*	145·1	152·8	158·3	162·5	165·6
19. Thomas		57·4*	87·2	106·5	116·6		127·0	133·7	
20. Int.		(3·1 I)	(1·6 R)	(2·0 I)	(1·2 R)		(1·0 R)	(1·4 R)	
21. Combination	4·8	57·4	86·8	103·9	115·2	(122·0)	(128·5)	(132·7)	(135·8)
22. Cu. Compton & Thomas	37·5*	92·2	117·8	†	148·4*	157·1	164·6		168·3
23. Int.	(2·8 R)	(1·6 R)	(1·5 R)		(1·4 R)	(1·4 R)	(1·6 R)		(1·1 R)
24. Calc.	37·5*	90·6	119·4	136·8	148·2*	155·9	161·4	165·6	168·7
25. Thomas		62·0*	86·7	108·2	117·8	127·1	132·9 ?		140·6
26. Int.		(1·8 R)	(1·2 R)	(1·8 R)	(1·5 R)	(1·4 R)	(0·9R)		(1·5 R
27. Combination	5·3	62·0*	87·6	(106·4)	118·2	126·9	(131·2)	(135·4)	138·1

† This might be covered by 132·9 ? (0·9 R) + 140·6 (1·5 R).

the first break is the strongest and there is a gradual falling off in intensity as n increases. Also, the breaks are all of the same character. This really looks like coming from a set of proper values.

We have also examined Thomas's data for cobalt, nickel and copper and Compton and Thomas's data for copper for similar relations. The results for cobalt are given in rows 10, 11 and 12 of Table IV, those for nickel in rows 16, 17 and 18, and those for copper in rows 22, 23 and 24. In each case the values used for determining A and b are indicated by asterisks. In the case of both cobalt and nickel the agreement is good; the first break to fail is that for $n = 9$, and there is the same graded falling off in intensity as was found for iron. Also, all the breaks are of the same character. In the case of copper the agreement is poor and the break for $n = 7$ fails. However, this failure may be only apparent, not real, and the disagreement is not bad enough to rule out the reality of the series absolutely. It is not much, if at all, worse than the disagreement between the data in the overlapping part of the tables given by Thomas and by Compton and Thomas for copper. The evidence for the reality of the series is indecisive in the case of copper and, perhaps, unfavourable; it is evidently much stronger in the case of iron, cobalt and nickel.

The values of the constants A and b in the equation V (volts) $= \mathrm{A} - b/n^2$ are :—For A, iron 180·0, cobalt 179·5, nickel 181·9, copper ? 185·0; for b, iron 2420, cobalt 2357, nickel 2357 and copper ? 2360. They evidently differ very little for these different elements.

It is not impossible that critical potentials occur for the number $n = 3$, for which the equation above yields negative values of V, *i.e.*, for transitions in which the term A would be the initial and $b/9$ the final levels. The calculated values and the nearest breaks are :—For iron, $n = 3$, $\mathrm{V} = -80·8$, Thomas 82·7 (2·3 R), Kurth 81·7, Andrewes, Davies and Horton 78; for iron, $n = 2$, $\mathrm{V} = -425$; for cobalt $n = 3$, $\mathrm{V} = -82·4$, Thomas 81·0 (1·2 R), A., D. and H. 83; for cobalt $n = 2$, $\mathrm{V} = -409·8$; for nickel $n = 3$, $\mathrm{V} = -80·0$, Thomas 79·8 (1·7 R), R. and C. 80·0 (3 to 1); for nickel $n = 2$, $\mathrm{V} = -407·4$; for copper ? $n = 3$, $\mathrm{V} = -77·2$, Compton and Thomas 73·5 (1·8 R), 82·5 (1·2 R), A., D. and H., 75; for copper ? $n = 2$, $\mathrm{V} = -405·0$. In the case of iron, $n = 3$, the calculated inflection is very near the value for $n = 5$ (83·2), which may account for the poor agreement, the breaks not having been resolved. There seems definite evidence of a break at the calculated point for $n = 3$ for both cobalt and nickel, but Compton and Thomas's data are definitely against it for copper. There is no evidence in the soft X-ray data of the existence of anything near $n = 2$ for any of these elements.

According to Rollefson the 180 volts (A) level is the M_5 ($= M_I$) level for iron, and his breaks at 51·7 and 100·3 volts are due to transitions between the other levels of 82·7 (2·3 R) ($n = 5$) and 131·8 (1·9 R) ($n = 7$) respectively and the $M_{3,4}$ ($= M_{II\ III}$) respectively. If this is the case a number of other combinations between the iron breaks should occur. If we take the break given by Thomas as 54·6 (3·1 I) to be $M_3 - b$ (5) and 103·5 (2·0 I) to be $M_3 - b$ (7), we have

$$M_5 - b\,(5) - (M_3 - b\,(5)) = M_5 - M_3 = 82{\cdot}7 - 54{\cdot}6 = 28{\cdot}1 \text{ volts and}$$

$$M_5 - b\,(7) - (M_3 - b\,(7)) = M_5 - M_3 = 131{\cdot}8 - 103{\cdot}5 = 28{\cdot}3 \text{ volts.}$$

Taking the mean of these, 28·2, to be the value of $M_5 - M_3$, we calculate from the values of $M_5 - b$ (n) in row 4 of Table IV the values of $M_3 - b$ (n) in row 9. The nearest lines of Thomas are given in row 7 and the intensity and character in row 8. The agreement is good apart from $n = 6$, 9 and 10. If these interpretations are correct, 82·7 (2·3 R) has to do duty for $M_5 - b$ (5), $M_5 - b$ (3) and $M_3 - b$ (6), for which the calculated values are 83·2, 80·8 and 84·5 respectively. The average of these numbers is 82·8, so that 82·7 is about where the break would be expected if these are unresolved. Again, although $n = 9$ and 10 are missing, there is a break 125·7 (1·5 R) almost half-way between them.

If the series already proposed for cobalt and nickel and, somewhat doubtfully, for copper are real, we should expect similar combinations to hold for them as for the iron breaks. Taking the breaks 55·8 (2·9 I) as $M_3 - b$ (5) for cobalt, and 57·4 (3·1 I) as $M_3 - b$ (5) for nickel, these breaks having a similar intensity and character to the break taken as $M_3 - b$ (5) for iron, and also taking 62·0 (1·8 R) as $M_3 - b$ (5) for copper, the $M_3 - b$ (n) breaks calculated from the $M_5 - b$ (n) breaks by the combination principle are given in rows 15, 21 and 27 of Table IV for cobalt, nickel and copper respectively. In rows 13, 14, 19, 20, 25 and 26 are given the corresponding nearest breaks followed by intensity data of Thomas or Compton and Thomas. In each case breaks are found at the calculated places up to $n = 8$ even for copper. (The voltages for $n = 4$ are too low in all cases for breaks to be detected by the photo-electric method.) The agreement is good except for $n = 7$, where the voltage is too high and the intensity too great in every case. It is difficult to know whether to consider this an objection to the reality of these combination series or not; it does not, at any rate, seem so serious an objection as if the discrepancy were scattered irregularly over various quantum numbers. It appears to us an argument in favour of the reality of these classifications that the agreement is substantial, and that there are no missing breaks in the lower quantum numbers, where the breaks would be expected to be more definite and stronger, as, in fact, they are. A good many of the combination breaks are for practical purposes coincident with the breaks in the other series. These cases are indicated by numbers underlined. It is interesting to observe that Thomas remarks that the iron break at 34·3 (2·0 R) "is not in good agreement with the corresponding breaks for cobalt and nickel. Careful check runs failed to detect any error in the determination of this voltage, so that the peculiarity seems to be real." This discrepancy is accounted for on the proposed classification which makes the corresponding iron break not 34·3 (2·0 R) but 28·8 (2·5 R). Our corresponding lines for iron, cobalt and nickel agree with his from $n = 5$ to $n = 8$ inclusive, but there are deviations for higher values of n as well as for $n = 4$.

If there were a satisfactory theoretical interpretation of this classification, the experimental evidence in favour of it would be regarded as strong. The real difficulty lies in the high value of the constant b, which has very much the same value in each case, viz., 2420 volts for iron and 2357 volts for cobalt and nickel. 2357 volts is equivalent to a nuclear charge of $13 \cdot 2e$. This high value is due to the wide spacing of the series breaks and the rather high quantum numbers which are required to co-ordinate them. Internal virtual orbits of this kind have not so far been contemplated in Bohr's theory of the structure of atoms, but there seems to be no reason why they should be absolutely prohibited, at any rate as temporary structures in excited atoms. The values of b are about what would be expected if the virtual orbits were executed around the completed L shell and the definite absence of the quantum number 2 seems significant in this connection. The fact that the wide spacing of these and similar soft X-ray excitation breaks prevents their explanation in any general way as fine structure, of the type usually considered, for accepted X-ray levels, has been pointed out already by Stoner.*

It seems to us that a large number of otherwise unco-ordinated facts can be brought into order if it is assumed that in the case of iron, cobalt and nickel, at any rate, it is possible for an electron to function as a series electron in relation to the structure bounded by the completed L shell without being very profoundly affected by the surrounding electrons. Whatever may be thought of this it is clear that some considerable departure from accepted views is required. There is evidence of this necessity not only in the soft X-ray field but in that of ordinary X-rays also. Nishina,* in examining the L absorption edges of the rare earths, finds two white lines on the short-wave side of this edge and distant from it about 32–34 and 89 volts respectively. This spacing is much too wide for any ordinary type of fine structure, and Stoner (*loc. cit.*) is content to leave it as an unsolved problem in X-ray theory. It seems, however, to be a difficulty of much the same character as that presented by the soft X-ray discontinuities. Furthermore, the process of atom building for the iron, cobalt, nickel group is very similar to that for the rare earths ; so that similar phenomena might be expected in both.

---

* 'Phil. Mag.,' vol. 2, p. 97 (1926).

On this scheme it would seem natural to assign the A level to M_5 $(= M_I)$ and the lower corresponding level of the combination series to $M_{3,4}$ $(= M_{II} M_{III})$. If that were the case these levels would have practically the same values for each element of the group under consideration, viz., for M_5, iron 180·0, cobalt 179·5; nickel 181·9 and ? copper 185·0 ; for $M_{3,4}$, iron 151·8, cobalt 151·0, nickel 152·1 and ? copper 154·2 volts. These values do not agree with those which have been deduced from ordinary X-ray data by Bohr and Coster,† Walter‡ and Stoner. Thus Stoner (*loc. cit.*, p. 102) gives for M_5, iron 93, cobalt 99, nickel 107, copper 120; and for $M_{3,4}$, iron 54, cobalt 58, nickel 69, and copper 78 volts. All calculated values of these individual levels depend either on the position of the K absorption edge or on the value of the K_γ frequency. Both of these are very difficult to measure accurately. At the same time it is unlikely, even if it cannot be said to be impossible, that the uncertainty is sufficient to bridge the disagreement. On the other hand, if there is all this complexity in the soft X-ray spectra of these elements, it may well be that the finer details of their ordinary X-ray spectra have been misinterpreted. Nishina's results for the rare earths show that there are several levels which remain practically constant in depth over a considerable change of atomic number for those elements. The difference between the A levels and the corresponding combination levels is about 3/4 of that between M_5 and $M_{3,4}$ calculated from the X-ray data. It is probable that this difference given by the X-ray data is not far out, as these levels form a pair of screening doublet levels. Also, it does not involve the difficult K absorption edge or the inaccurate K_γ line required to get either level separately. It does, however, involve the difference between an L frequency and the difference between two K frequencies so that a discrepancy of about 10 volts hardly seems unreasonable. In any event there are a considerable number of soft X-ray inflections left unarranged, so that perhaps it is premature to try to settle the nature of these levels at this stage.

---

* 'Phil. Mag.,' vol. 49 p. 528 (1925).

† 'Z. f. Physik,' vol. 12, p. 342 (1923).

‡ 'Z. f. Physik,' vol. 30, p. 357 (1925).

3. *Experiments with Deposited Metals.*

A way of testing whether the discontinuities which we have been investigating are instrumental or are real properties of the materials under investigation is to examine the emission from a target made of a given material as successive layers of some other substance are gradually deposited on it. Under these circumstances we should expect the discontinuities characteristic of the first substance to fade away and those characteristic of the second substance gradually to appear if the discontinuities are real properties of the materials, but not otherwise. This question is also of importance from the point of view of the necessity of ascertaining the effects which are likely to be produced by the deposition of the matter of the cathode on the targets, which takes place in all experiments of this kind.

For the original target we chose carbon, which, apart from some fine structure, has at most four strong discontinuities between about 30 and 360 volts, viz., at 39·5 (7), 81·2 (7), 265·0 (2) and 292·8 (10). In our former paper we expressed an opinion that it was not certain that the effects at 81·2 and 265 volts were not due to impurities. The present work tends to increase this suspicion, and also as regards 81·2 so does the work of Compton and Thomas.* They give a list of 62 discontinuities for carbon between 8·95 and 155·1 volts, and the corresponding break, which they measure as 78·70 (4) only appears on one out of six curves, although it is one of the strongest in their list and all the rest appear on at least half the curves. However, they get a fairly strong break (strength 3·1) at 73·54, so that the matter may not be so simple.

As the metal to be deposited we chose tungsten, as we have to use a hot tungsten filament as a source of electrons ; so that this element will gradually get deposited on the target in any event.

We used a revolving target of the type already described, except that the rod below, which supported the target, and the rod at the top, which supported a nickel mirror to enable the target to be aligned, were cut off from the same carbon block as the target itself. The object of this was to avoid any risk of obtaining the emission from any substance other than carbon. The edges of the back face of the target were bevelled to distinguish it from the front. The lower rod fitted into a cup cut in a rod of nickel. In all, seven series of experiments were made. In the first three the rod was held fast by melting copper or silver round its lower extremity. In the latter series it was found possible to wedge the rod firmly in the nickel cup without the addition of any other substance.

---

* 'Phys. Rev.,' vol. 28, p. 601 (1926).

Our original plan was to clean up the target and the neighbouring parts of the tube generally by bombarding the bevelled face of the target with a high potential electron current from the hot cathode; afterwards to turn the other face of the target towards the cathode and then to make the usual photo-electric tests of the generated X-rays as a function of the applied voltage, from time to time, as the target got covered with tungsten by thermal evaporation from the hot cathode filament. However, we did not get this plan to work with complete certainty in practice because it does not seem possible to restrict the deposition of foreign matter during bombardment to the side of the target which faces the cathode. In order to maintain good vacuum conditions in an extended experiment of this kind it is necessary, or at least helpful, to have a new bombardment at frequent intervals. Moreover, it does not seem possible to deposit the tungsten by thermal evaporation, except very slowly, at temperatures which do not endanger the life of the filament. For these reasons, by the time we got a deposit on the target we were uncertain whether it got there as a result of evaporation or of bombardment. The precise manner in which the deposits were formed must therefore to this extent be considered as doubtful.

The filaments used have been short. Their current-carrying capacity is rated at 3·87 amps., their mass per unit length being 0·0039 gm. The total potential fall down the filament has rarely been greater than 3 volts.

In the two first series of experiments and also in the last series the photo-electric current was measured by the steady deflection method, using the alcohol-xylol and lampblack resistance already described. In the intervening series the following compensating method was usually employed :—

The scheme is very similar to that used by Compton and Thomas, but it was planned independently. It was, in fact, intended for the preceding experiments with iron, but owing to a long delay in the delivery of the high resistance potentiometer, those experiments were completed before it could be put into operation. A diagram of the connections with the apparatus as it was actually first used is shown in fig. 2. The principle of the method will be clearer by reference to fig. 3. i_p is the photo-electric current to the electrometer, r_1 and r_2 are high resistances through which currents i_1 and i_2 pass respectively.

Actually r_1 was the xylol-alcohol and r_2 a lampblack resistance. A voltage V is applied across the resistances r_1 and r_2, whilst i_p flows in at the junction between them. In effect we thus diminish the deflection which i_p produces on the electrometer by the superposition of another current of opposite sign. Let v be the steady potential at the electrometer. Then

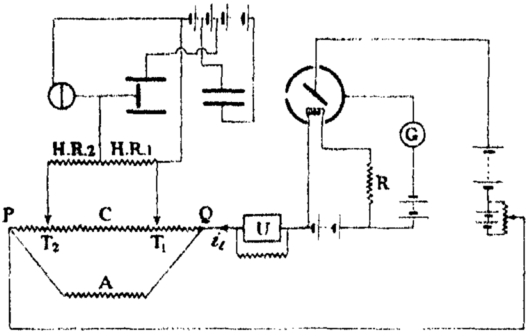


FIG. 2.

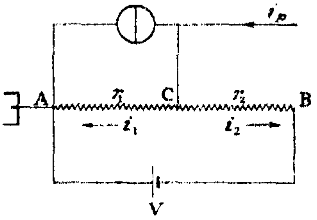


FIG. 3.

$$v = i_1 r_1,\quad v - \mathrm{V} = i_2 r_2$$

and

$$i_p = i_1 + i_2 = v/r_1 + \frac{v - \mathrm{V}}{r_2} = v\,(1/r_1 + 1/r_2) - \mathrm{V}/r_2.$$

Thus

$$v = (i_p + \mathrm{V}/r_2)\frac{r_1 r_2}{r_1 + r_2},$$

and we have no deflection when $\mathrm{V} = -i_p r_2$.

Also, if C is the capacity of the electrometer and its connections,

$$\mathrm{C}\,dv/dt = i_p - i_1 - i_2 = -v\,(1/r_1 + 1/r_2) + i_p + \mathrm{V}/r_2,$$

and putting $\alpha = -(1/r_1 + 1/r_2)$, $\beta = i_p + \mathrm{V}/r_2$ and integrating,

$$\frac{\mathrm{C}}{\alpha}\log\left(v + \frac{\beta}{\alpha}\right) = t + c_1.$$

If $v = 0$ when $t = 0$, $c_1 = \frac{\mathrm{C}}{\alpha}\log\frac{\beta}{\alpha}$; so that

$$v = \frac{\beta}{\alpha}\left(e^{\frac{\alpha t}{\mathrm{C}}} - 1\right) = (i_p + \mathrm{V}/r_2)\frac{r_1 r_2}{r_1 + r_2}\left(1 - e^{-\frac{t(r_1 + r_2)}{\mathrm{C} r_1 r_2}}\right).$$

So far this is the method as Compton and Thomas used it. It seemed convenient, however, to use the thermionic current to supply the voltage V. To accomplish this we inserted a high resistance potentiometer (C, fig. 2) and allowed the thermionic current i_t to flow through it. The potential was then tapped off at the two points T_1 T_2 as shown in fig. 2. If r_3 is the total resistance of the potentiometer and F is the fraction of this which is tapped off, the voltage V used is $= -\mathrm{F} . r_3 i_t$.

Thus for zero deflection of the electrometer we have

$$i_p = -\mathrm{V}/r_2 = \mathrm{F}\frac{r_3}{r_2}i_t \quad \text{or} \quad i_p/i_t = \mathrm{F} . r_3/r_2.$$

In this way we can measure i_p/i_t directly from the potentiometer readings. Since i_p is approximately proportional to i_t, small fluctuations in the thermionic current should produce no unsteadiness. There are no deflections to be measured by this method, the only moving instrument being the electrometer, which is only used as an indicator. The thermionic current may be increased up to the limit of the current-carrying capacity of the potentiometer, which enables a further increase in the accuracy of the method to be obtained. Owing to the lack of some auxiliary apparatus we have not been able to exploit to the full the accuracy to be expected from this method in the present work. But, for the problem under consideration, it has not been necessary to do so.

A disadvantage of the method lies in the fact that we can no longer use a voltmeter directly across the tube, as we should then add the current through the voltmeter to the thermionic current, and this would alter the shape of the curves. We can, however, measure the potential drop by putting a voltmeter between the target and the end P (fig. 2) of the potentiometer and then measure with a separate voltmeter the potential drop (up to 15 volts) across the potentiometer PQ. The potentiometer had a resistance of 15,000 volts. In the earlier work with this method the only voltmeter available for measuring the drop in the potentiometer had a resistance of only 1100 ohms. This reduced the effective resistance of the potentiometer to such an extent that it was impossible to use it above about 200 volts. Above this range at this stage the thermionic currents were measured with the unipivot galvanometer U. This method is not accurate enough and the numbers obtained in this way have not been used. There is also a risk of trouble due to leakage currents to the line from the electrometer to T_1. Later on, a voltmeter with a resistance of 3000 ohms was used (in series 4, 5 and 6), and with this the whole range of voltage required could be covered.

In the course of this work there has been overlapping, from time to time, of measurements by the following three methods:—(1) the usual constant deflection method, (2) the compensation method with voltmeter correction, and (3) the compensation method with galvanometer correction. All three methods give identical results so far as we have been able to ascertain. The compensating method would be ideally worked with some static device for measuring the voltage, which could be put directly across from the filament to the target. With the electrometer in use (sensitivity 3000 mm. per volt) and working at maximum sensitivity with the potentiometer, a change of 1 per cent. in i_p will give a deflection of about 6 cm.

A curious effect was observed on one occasion. After bombardment the front face of the target was found to be covered with a displaced and distorted image of itself. We have endeavoured to draw the appearance of it in fig. 4. The shaded part was a dark brownish black, whilst the unshaded portions were greyish in colour. The mirror support above had the normal carbon colour. Possibly the explanation of these appearances is as follows:—The bombarded target is near the axis of a surrounding nickel cylinder and positive ions will be emitted from it and flow along the lines of the field of force towards the cylinder. There are, however, a large number of electrons in the field. These will first neutralise the positive ions and then charge them negatively so that they will be pulled back along the lines of force to the target. If the target were a line along the axis of the cylinder, they would come back exactly to the point from which they started, but as it is an extended object they will form a distorted and displaced image of it.

FIG. 4.

Of the seven series of experiments it will perhaps be sufficient to consider series I in detail. This was one of the shortest, the range from 250 to 350 volts only being examined for critical potentials. No attempt was made to produce impurity effects by evaporation, bombardment only being relied on. The operation of this method is uncertain, as it sometimes cleans up the target instead of depositing foreign matter on it. The results of the various operations are collected in Table V.

Table V.—Series I.

Experiment No.	Details.	Critical Voltages.			
1	The bombarded (back) face		absent	297·2	absent
2	The clean (front) face		absent	296·2	absent
3	Fresh bombardment of front face (blue glow)		287·7	?	335·7
4	Fresh bombardment of front face....	271·7	286·7	297·2	
5	Fresh bombardment of front face			297·8	329·6
6	Fresh bombardment of front face (blue glow)			293·4	absent
7	Clean back face			296·8	
	Mean values	271·7	287·2	296·4	332·6

It will be seen that the initial bombardment did not deposit any appreciable quantity of foreign matter on the target, only the carbon K inflection at 296·4 volts being found both on the bombarded back face and on the front face after this bombardment. Successive bombardments then gave rise to the deposition of increasing amounts of matter having discontinuities at about 270, 287 and 332 volts. The gradual growth is shown in fig. 5, which is a plot of the results of experiments Nos. 2, 3 and 5. In Nos. 5 and 6 the points are taken far apart so that the resolving power is low. In No. 3 the carbon inflection has almost if not entirely disappeared. This can hardly be due to the carbon having got entirely covered up, because the impurity breaks are very weak at this stage, and in No. 5, when the impurity break at 332 has become quite strong, the carbon break is strong also. It is quite strong in No. 6 as well. Also, it cannot be due to the position of the carbon inflection having shifted to 287, because both inflections are visible on No. 4. It is probably an illusion arising from the presence of two inflections close together. If enough experimental points are not taken, the points can then run so as to make it appear as if an inflection were missing.

There is no need to consider in similar detail all the seven series of experiments. Perhaps the quickest way of proceeding with the matter is to consider a list of the inflections which were obtained in the various series. This is given in Table VI. All the values found are included except a few which were only got once, which are probably due to experimental errors or to accidental impurities, and in any event have no particular significance. The mean values are given in the row below that for series VII and in the two next rows are given the values we formerly obtained for W and C respectively. It is at once evident that we have succeeded in getting in addition to the four carbon inflections every one of the old W inflections in the range 35 to 350 volts covered by the experiments. The only values about which there can be any question are the weak W inflection 40·3 (0), which coincides with carbon L, and the carbon inflection 265·0 (2), which is indistinguishable from W 263·7 (4). It is doubtful whether this really is a carbon inflection at all. It may have been due to W contamination in the former experiments. The carbon L break was examined at the beginning of series III and measured at 39·8 volts. It subsequently moved to lower values owing to being mixed with W 36·8. On one curve the two were found present together and resolved. Except for inflections which appeared in only one series of experiments and which may have been caused by accidental contaminants, the only new ones are those at 226, 284·1 and 288·4 volts. The inflection at 331 is 5 volts below

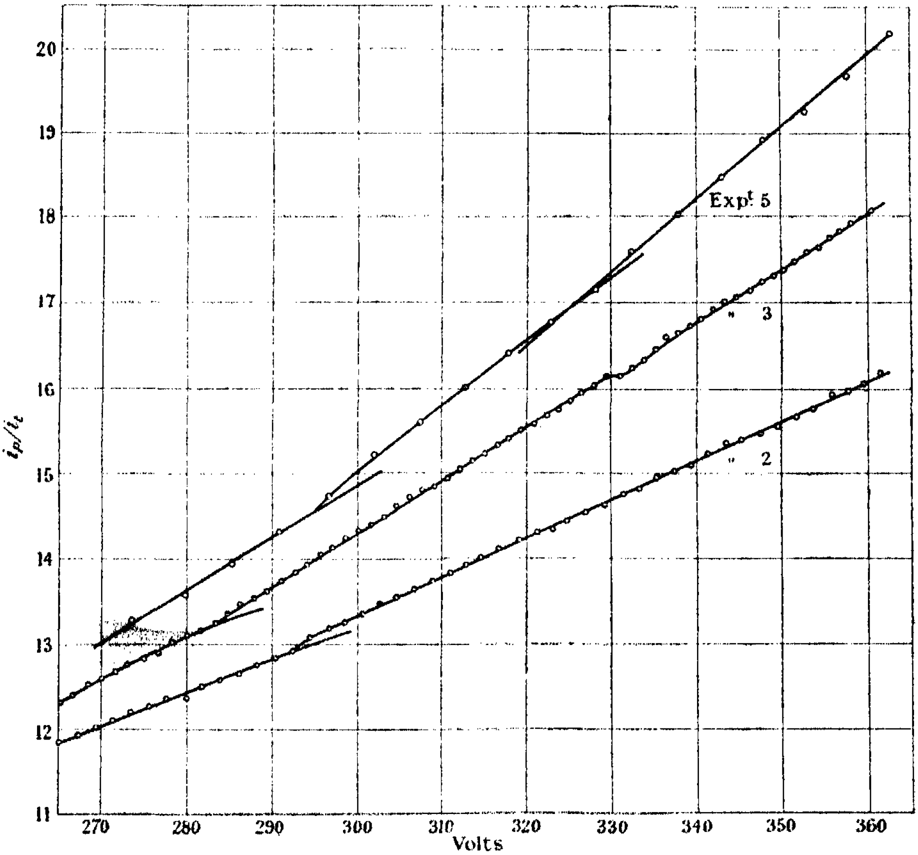


FIG. 5.—Experiment 3 is displaced 1 upwards and experiment 5 is displaced 0·5 upwards, each with respect to experiment 2.

Table VI.—List of Critical Voltages found in the Seven Series of Experiments.

Series No.																				
I	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	(271·7)	—	287·2	296·4	332·6
II	—	—	—	—	—	—	—	—	—	—	—	—	—	225·9	—	—	—	288·4	296·0	334·6
III	37·9	44·6	54·0	75·6	—	91·4	—	128·1	142·6	152·4	—	—	200·0	228·4	—	—	284·3	287·7	297·0	334·0
IV	—	—	—	—	80·0	—	104·0	130·7	—	153·2	166·8	182·6	197·5	226·4	—	262·5	283·9	290·3	300·8	328·0
V	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	284·0	?	(293·0)	329·0
VI	—	—	—	70·2	82·9	—	104·3	123·6	—	155·3	—	—	193·5	—	242·0	262·6	284·0	?	295·9	329·5
VII	36·8	46·6	—	70·6	—	—	108·2	—	—	—	—	—	—	—	248·4	267·4	—	—	295·0	—
Means	37·4	45·6	—	72·1	81·5	—	105·5	127·5	—	153·6	—	—	197·0	226·8	245·2	264·2	284·1	288·4	296·3	331·3
W. (Int.)	36·8 (5) 40·3(0)	44·9 (2)	—	70·8 (7)	—	—	108·2 (0)	127·2 (2)	—	150·0 (1)	—	—	191·0 (8 to 4)	—	245·2 (4)	263·7 (4)	—	—	...	336·0 (10to5)
C.	39·5(7)		—	—	81·2 (7)	—	—	—	—	—	—	—	—	—	—	265·0 (2)	—	—	292·8 (10)	—
C.X.R.L.	L		—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	K	—
W.B. & C.	33·9 N_{VI} N_{VII}	39·3 O_{II} O_{III}	—	70·4 O_I	—	—	—	—	—	—	—	—	—	—	241·0 N_V	253·5 N_{IV}	—	—	—	—
W.N.	37·9 N_{VII}	39·3 N_{VI}	—	70·4 O_{II} O_{III}	77·2 O_I	—	—	—	—	—	—	—	—	—	249·0 N_V	255·0 N_{IV}	—	—	—	—
Ni R. & C.	38·6 (3/0)	—	—	—	80·0 (3/1)	88·9 (6)	99·0 (6)	—	140·3 (4)	—	—	—	—	—	245·5 (7)	—	282·0 (1)	288·0 (8)	316·4 (7)	335·0 (7)
Ni. Th.	—	44·7 (1·5 I)	—	—	79·8 (1·7 R)	91·2 ? (0·9 R)	106·5 (2·0 I)	127·0 (1·0 R)	—	156·8 (1·4 R)	167·3 (1·0 R)	178·2 186·2	197·7 (1·8 R)	217·7 228·2	—	265·9	280·0	286·5	—	—

the former value. In some experiments this inflection was observed to be double. This feature may account for the discrepancy and also for the variation (7·6 volts) between the values found in the different series. The carbon K value was found consistently about 4 volts higher than in the former experiments. Even if the unreliable values 300·8 and 293·0 are thrown out, the average will still be 3·2 volts higher at 296·0 and the other determinations are very consistent with each other. This may be an effect of the presence of a layer of tungsten or other foreign matter on the carbon, but our observations do not afford any evidence in favour of this view. What seems more likely is that in the old experiments in which the resolution was less, some of the effects at 284·3 and 288·8 were present and influenced the results. The point at 331 may have been present in those experiments as no detailed examination was made at high enough voltages to discover it with certainty.

A study of the data made available in the first five series of experiments showed that the various deposited inflections were not equally interdependent, but that they should be classified in groups such that the members of each group appeared and disappeared, or strengthened and weakened, together independently of the other groups. For example, it was noticed that the strong inflection at 331 and the doublet at 286 were easily produced, that they could subsequently be got rid of or at least very much reduced by bombarding the target, and that they appeared and disappeared together. It became clear from series III that the inflections at 197, 153 and 127 behaved in the same way as these. Let us call these inflections group I. In series IV two strong inflections appeared near 265 and 108 as well as some fainter ones which were not found in series III. In series VI we did not start with a fresh piece of tungsten filament but purposely used the same piece of filament as had been used in series V. The object of this was to test the idea that the group I inflections were due to some more volatile contaminant of the tungsten filament. In this series the inflection at 331 was not found at first, 286 was doubtful, and the group I inflections were weak generally. On the other hand, 265 was strong and was accompanied by 245, 108 and 71. We shall refer to these inflections as group II. The weak inflections at 183 and 167, which appeared in series IV, were not found in series VI or, in fact, in any other series. This suggests that they do not belong either to group II or to group I. There is, however, evidence that they belong to each other. Although the inflection 226 appeared in series II as well as 331 and 288, they were not present at the same time. This shows that it does not belong to group I. Its absence from series VI suggests that it does not belong to group II either. There is some evidence, not very convincing, that 91·4 is associated with 226.

In this way we have arranged the inflections in the following scheme of related groups :—

Table VII.—Groups of Critical Voltages.

Group I	81·5	127·5	153·6	197·0	284·1	288·4	331·3 (double)
Group II	37·4	45·6	72·1	105·5	245·2	264·2	
Group III	166·8	182·6					
Group IV		91·4 ?	226·8				

We come now to the consideration of the origin of these different groups. There can be little doubt that group II is due to tungsten. In the row in Table VI marked C.X.R.L. are indicated the carbon X-ray levels and in the two succeeding rows are given respectively Bohr and Coster's* and Nishina's† values for the X-ray levels of tungsten. It will be seen that two of the levels in group II coincide with Nishina's N_{VII} and $O_{II\,III}$ and three others are very close to N_{VI}, N_V and N_{IV}. We are not sure whether the O_I level is known with enough certainty to make the gap between 107 and $O_I = 77$ impossible. The estimated location of some of these relatively superficial X-ray levels has moved a good deal recently, as can be seen by a comparison of Bohr and Coster's and Nishina's values. Apart from this discrepancy, there are the same number of strong inflections as of X-ray levels in the region investigated, and they are practically in the positions where the application of the combination principle to X-ray measurements has put them. In any event the 107 level may not be a W level at all. It does not appear in some unpublished W data of Richardson and Roy. Furthermore, there should not be an exact agreement, as Nishina's data were got with tungsten oxide WO_3, not with the metal.

If the assignation of the group II inflections to tungsten is correct, it follows that the majority of the inflections found must be attributed to contaminants. Now in all the first six series of experiments the same reel of tungsten wire had been used. Accordingly, in series VII we decided to try the substitution of tungsten wire from a different reel. In all, ten experiments were made in this series, covering the whole range a number of times. Except for a faint indication of 331 in the first and third experiments, the group I inflections were entirely absent, the group II inflections were fully developed, and each one was present in every curve extending over the range in which it occurred, and the only inflections present not belonging to group II was something near 220 and this had disappeared at the end of the experiment. We regard the results of this series as conclusive that the real tungsten inflections are those in group II.

---

* *Loc. cit.* Values given by Thoraeus 'Z. f. Physik,' vol. 26, p. 398 (1924), are identical with these within the limits of error.

† *Loc. cit.*

It seemed likely from these experiments that the group I inflections were caused by some more volatile contaminant which was present in the tungsten filament used in the first six series of experiments and absent from the different sample used in the seventh. There is also evidence from the old tungsten experiments that some of the group I inflections were due to contaminants. Thus the 331 and 196 inflections were distinguished from all the others by being assigned very high but variable intensities, viz., 336 (10 to 5) and 191 (8 to 4) respectively. It is quite possible that the 286 doublet was present in the old experiments and was overlooked because it so happened that all our more accurate examinations of this region either started or finished at 290 volts and inflections so close to the end of a run are often not very obvious.

We were strongly inclined to attribute the group I inflections, or at any rate most of them, to thorium. But we have had the sample of wire used in the first six series of experiments analysed and there is no trace of thorium to be found in it.* We are also informed that it is not very likely to contain any other impurity. Accordingly we are disposed to attribute this group of critical potentials to nickel, which was present in some form or other in the bombarded parts of the apparatus used on each occasion when these breaks appear. That most of them coincide with the strong nickel inflections is clear from the values taken from our former paper in the bottom row but one of Table VI. The weaker ones agree with values given by Thomas, some of which are quoted in the bottom row of Table VI.

Since the foregoing was written we have found confirmation of the view that the group I inflections are not caused by tungsten but by some adventitious impurity from an unexpected and independent source. A paper by H. E. Krefft in the last number of the 'Annalen der Physik' (vol. 84, p. 639) contains an exhaustive examination of the secondary electron emission from tungsten which has been heated to high temperatures. *When the results are plotted with the proportion of secondary rays to primary rays as ordinate against the primary voltage as abscissæ he obtains, in contrast with previous investigators who had used cold tungsten, curves which show a number of inflections almost identical in position with those published in our previous paper. He states that although the other inflections remain, those at 390, 330, 190 and (140, 160) disappear on heating the tungsten to higher temperatures. Evidently 330 and 190 are the same as our 331 and 197. It seems likely that 160 is the same as our 167·5 and 140 the same as 142·6 mixed with 153·6, since Krefft's values correspond to our uncorrected, not our corrected, values. 153·6 may be too high, as our previous determination made it 150·0 after rejecting some low values. It appears from Krefft's values that the inflection 395·0 (2) of our former paper should also be added to group I or, possibly, to group III. As the next of our former critical voltages 419·0 (2) was close to the Bohr level N_{III}, we may now say that all the strong tungsten inflections between 35 and 425 lie fairly near the levels deduced from hard X-rays, and there are none which do not correspond to Bohr levels. This does not preclude the possibility of there being a fine structure superposed on these levels ; but according to the present interpretation the fine structure found by Compton and Thomas for tungsten in the neighbourhood of 197 volts is probably due not to tungsten but to something else. Krefft's apparatus also contained nickel parts which would get hot. This tends to support the attribution of the group I effects to nickel. However, the possibility does not seem to be absolutely excluded that the group I effects might be due to tungsten in some unusual physical condition.

---

* We are very much indebted to Mr. L. St. C. Broughall for arranging for this to be done for us.

As a result of the present investigation we are led to reconsider some of the critical potentials given in our former paper. In the case of carbon the weak inflections found at 265·0 (2) is almost certainly due to tungsten 263·7 (4). The strong inflection at 81·2 (7) has not always occurred in the present work, which suggests that it is due to some impurity (? Ni 80·0 (3/1)). This leaves the only certain inflections for carbon as 292·8 (10) and 39·5 (7), the K and L levels respectively. The L level was found to be complex, but this matter needs re-investigation on account of the existence of levels due to tungsten, nickel and iron, among other elements, in this neighbourhood. The only inflections left for tungsten are the complex affair analysed as 36·8 (5), 40·3 (0), 44·9 (2), (N_{VII} and N_{VI}), 70·8 (7), ($O_{II\,III}$), 108·2 (0), (? O_I), 245·2 (4), (N_V), 263·7 (4), (N_{IV}) and 419·0 (2), (? N_{III}). In the case of nickel we are led to put back 336 (10/5) and 395 (2) previously suspected to be due to tungsten. The inflections 38·6 (3/0), 80·0 (3/1) and 288 (8) suspected as possibly due to carbon are probably nickel inflections. In view of the enormous amount of fine structure exhibited by the metals of the iron group all we are inclined to claim for these nickel inflections is that they indicate approximately where the breaks will be found when the points are not taken very close together, *i.e.*, with rather low resolving power.

We are glad to have this opportunity of expressing our indebtedness to the Department of Scientific and Industrial Research for financial assistance to one of us (F. C. C.), without which the continuance of this research would have been impossible.

Intensity Measurements in the Secondary Spectrum of Hydrogen.

By Prof. L. S. ORNSTEIN, Dr. W. KAPUSCINSKI and Miss J. G. EYMERS.

(Communicated by O. W. Richardson, F.R.S.—Received January 20, 1928.)

Prof. McLennan and others have published in these Proceedings* a paper in which they have given results of their measurements of the intensities in the secondary spectrum of hydrogen. The experimental results mentioned in the tables of that paper, however, are only densities, and no attempt was made to find from those densities the real values of intensity in the spectrum. The relations between density and intensity, however, is not so simple as McLennan supposes, as is shown in fig. 1 which follows this. In this paper it will be shown

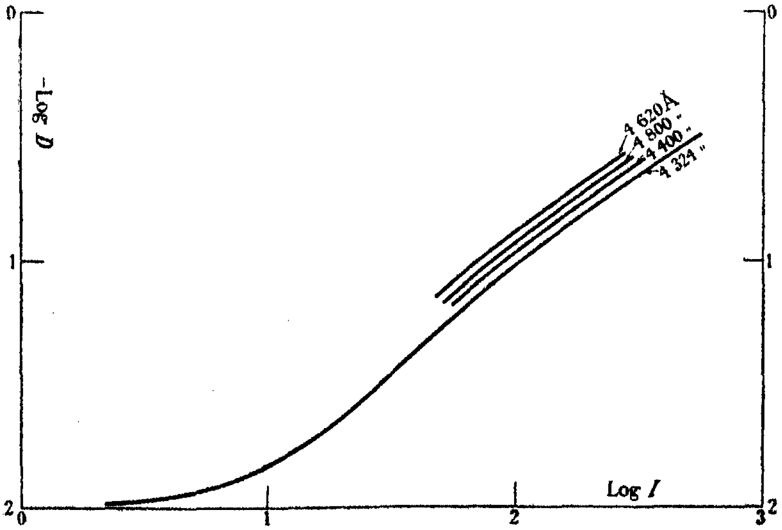


FIG. 1.—Density-Intensity Curves.

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 277 (1927).

that *for theoretical investigations* on the secondary spectrum it would be dangerous to use the figures given by McLennan. We shall do this at the moment for a region from 4500–4900 Å.; the real intensities in this region will be shown to differ considerably from those given by McLennan. The intensities of about 230 lines have been measured by the method developed in our Institute. Further, our work shows that the source of error, already mentioned by McLennan—*e.g.* (p. 281), the difference of sensitivity of photographic plates in this region—can be considerable. Fig. 2 represents the sensitivity as a function of wave-length for the plates used (Ilford Special Rapid). Measurements for other regions of the secondary spectrum will be published soon.

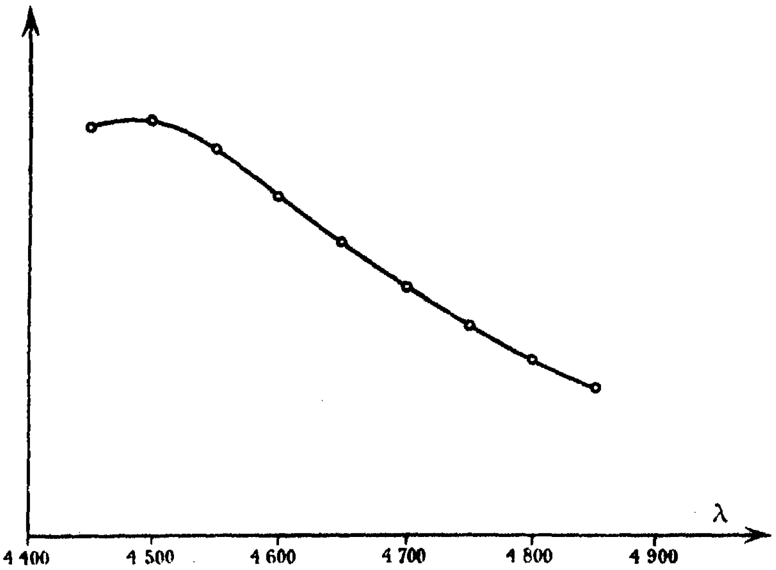


Fig. 2.—Sensitivity Curve of our Plate.

The application of the intensity rules is hardly possible at present, as the number of lines of which the classification is certain is too small, and also the wave-lengths are not known with sufficient accuracy. In this paper we give, in addition to the measurements of intensities, the wave-lengths of a number of complex lines not described or uncertain in existing literature.

Experimental Method.

As we have mentioned already the measurements of the intensities are accomplished according to the Utrecht method. The light-source was a discharge tube (fig. 3), consisting of two spherical vessels, joined by a capillary tube.

1 cm. long. The discharge tube was provided with an oxide-glow-kathode and the tube was operated at constant voltage; the voltage between the electrodes was about 600–800 volts. The hydrogen was admitted through a Palladium tube (P). For our definitive plates we have used a pressure of about 0·5 mm. and currents of 16 and 4 m. A.; the times of exposure amount to 3 and 4 hours; we should expect the proportionality factor of the intensities of the two series of measurements to be about three (see column 4 of Table I). The photographs were obtained in the first order with a concave Wood grating of 1 m. radius and 30,000 lines per inch, with Rowland mounting. The dispersion was 8·4 Å. per millimetre; the lines of 0·02 mm. distance on the plate were fairly separated in photometrical work; visually we could resolve lines of smaller separation.

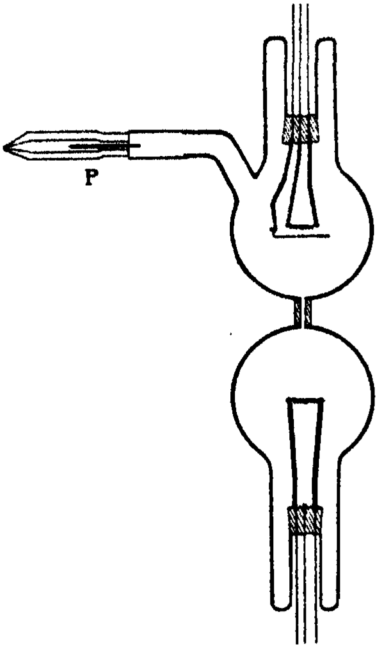


FIG. 3.

On the same plate with the hydrogen spectra were photographed 12 continuous spectra, as density marks, of a Wolfram lamp whose intensity distribution as a function of wave-length has been previously determined.

The various intensities were obtained by the aid of the variation of the width of the slit of the grating spectrograph. The width of the slit was chosen in such a way that at least one of the diffraction maxima at each side was completely within the opening of the beam; and therefore the diffraction phenomena have no influence on the proportionality between the intensity and the width of the slit.

The density in the spectra was measured by means of a Moll registering microphotometer. The density marks were measured at four properly chosen wave-lengths and photometered in a direction perpendicular to the spectrum. We obtained in this manner density-intensity curves at four different wave-lengths, which were almost parallel (fig. 1). In order to know the sensitivity of the plate for all wave-lengths considered, we have photometered two density marks in the direction of the spectrum. In the regions which we have investigated, 4500–4900 Å., the value of the sensitivity depended strongly on the wave-lengths, as we have mentioned before. Every line had to be referred to the corresponding density curve. The enlargement of the plate on the microphotogram was 20 or 80, the last being used in connection with the analysis of closely adjacent lines. The method used for this analysis will be explained now.

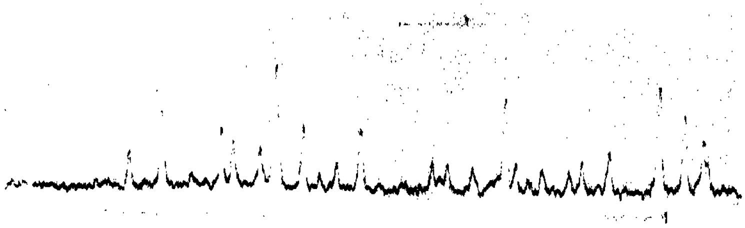


Fig. 4.—A Portion of a Microphotometer Curve (enlargement 1 : 20).

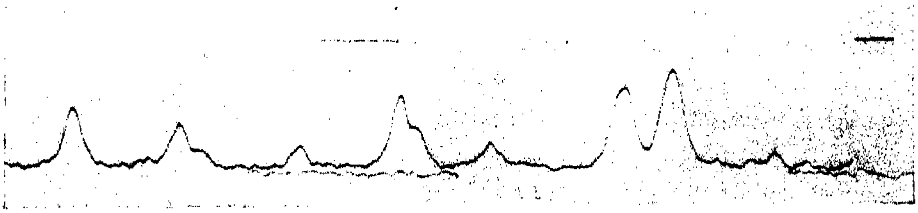


Fig. 5.—A Portion of a Microphotometer Curve (enlargement 1 : 81).

It is known that lines which are very near each other can have a great mutual influence, both as to wave-length and intensity, when the maxima are taken as a measure for distance or intensity. When both lines have about the same intensity, the influence is reciprocal, but when one of the lines is far stronger than the other, the wave-length and the intensity of the former is not influenced by the other line. Only the line of small intensity then is strongly deformed (figs. 6 and 7). In order to find the intensity and wave-length, undisturbed by this effect, of the components of complex lines, we have proceeded as follows :—

The portion of the microphotometer curve is enlarged by an episcope and the enlarged curve is measured, the intensity for 20 or 30 points being determined, and in this way the intensity curve is drawn. In fig. 7 one sees immediately the axis of symmetry of the strong line ; the undisturbed side then can be reflected at this axis (points given by ×) and the intensities obtained can be subtracted from the whole intensity (points given by +). The intensity which is left is that of the second line. It must be shown that this second line is also symmetrical about an axis and, besides, the width at the half intensity of

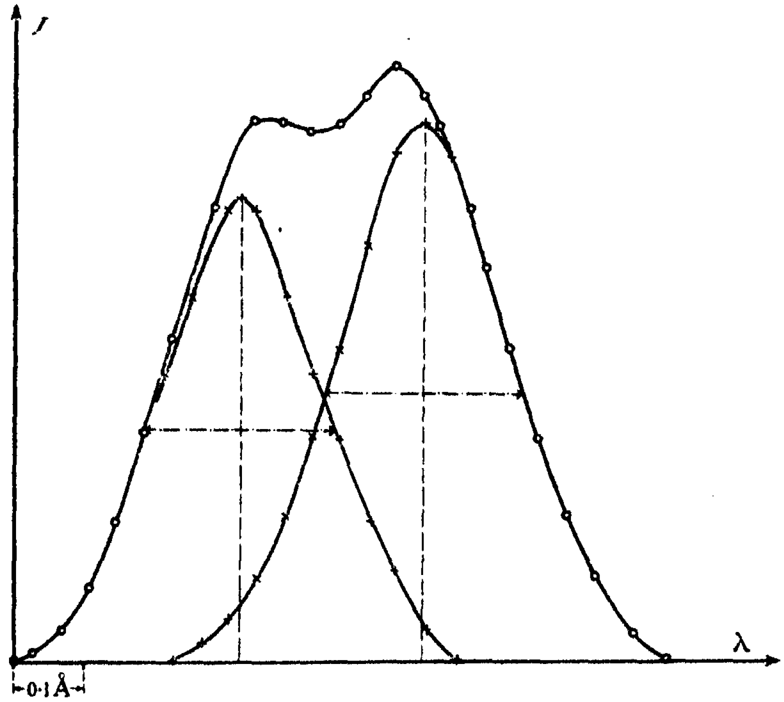


FIG. 6.—Analysis of a "Doublet."

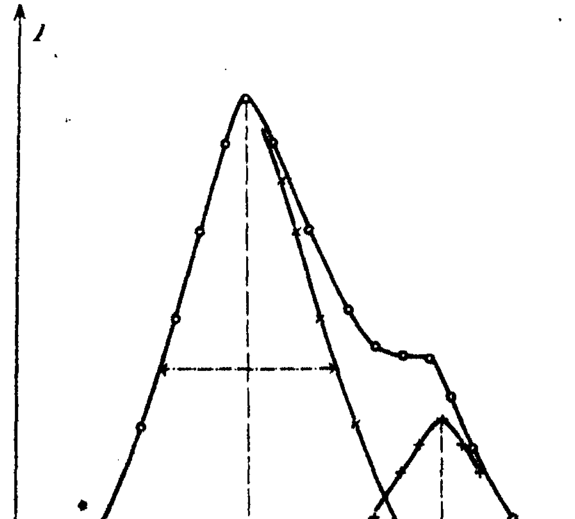


FIG. 7.—Analysis of a "Doublet."

both lines ought to be the same. When those conditions are not fulfilled, the analysis is not right. In a case shown by fig. 6 the analysis is much more difficult; the symmetry axis of both lines are unknown; the sum of the intensities of both lines must be the curve that is given; the value of the half-width of the lines and the fact that they are symmetrical about their axis is the proof of the exactness of the analysis,

In Table I one finds the results of our measurements; column 1 gives the wave-lengths, which are given by Merton and Barratt,† Tanaka‡ and Deodhar§; column 2 gives the intensities of the lines in the photograph with a current of 16 m. A., column 3 those in the photograph with a current of 4 m. A., and column 4 the proportion between those two series of measurements. The classification of the lines, known at this moment, is given following Richardson,‖ or Allen and Sandeman.¶ Indications with ? mean that the classification is uncertain.

D means "double," D ? "probably double," but we could not separate them.

Table I.

λ						λ					
4891·22	4·6	—	—		3α1Q(1)?	4811·67	2·9	—	—		II BeQ(4) 2α0Q(5)
88·44	2·7	—	—		—	07·33	8·6	2·8	3·1		II BdQ(1)
86·40	4·5	—	—	D ?	—	04·06	1·7	—	—	D ?	—
83·82	4·5	—	—		—	02·03	8·2	—	—		—
81·55	3·8	—	—		—	4799·00	2·0	—	—		—
78·10	3·7	—	—	D	—	97·76	32·6	12·0	2·7		II BeQ(5)
75·92	17·9	—	—		—	97·02	18·7	6·6	2·8	D ?	—
74·2*	5·3	—	—		—	95·14	3·3	—	—		2α0Q(4)?
73·05	47·7	—	—	D	—	94·22*	4·2	—	—		—
69·42	12·9	—	—		—	93·95	15·7	6·0	2·6		—
67·05	8·9	—	—		—	92·52	3·3	—	—		—
66·2	3·2	—	—		—	91·17*	1·9	—	—		—
56·54	37·9	—	—		—	90·82	7·4	—	—		—
49·89*	6·4	—	—		—	89·40	11·5	3·8	3·0	D ?	—
49·34	39·1	—	—		—	86·25*	5·0	—	—		—
43·69	5·9	—	—		—	86·04	7·3	2·5	2·9		II BcQ(1)
43·35	9·2	—	—		—	85·20*	3·3	—	—		—
42·45	6·2	—	—		—	84·85	5·4	—	—		II BcQ(2)
38·25	17·8	—	—		—	83·47	1·5	—	—		2α0Q(3)
32·79	16·3	—	—		—	82·85*	2·4	—	—	D ?	—
31·60	3·8	—	—		—	82·48	4·2	—	—		—
30·69	5·0	—	—		—	80·94	24·6	7·3	3·4	D ?	$4\beta_4P(3)$
24·58	6·8	—	—		—	77·42	15·8	5·2	3·0		—
17·51	5·7	—	—		—	74·29	2·7	—	—		2α0Q(2)
13·61	11·3	—	—		—	70·78*	6·0	—	—		—

---

† 'Phil. Trans.,' A, vol. 222, p. 369 (1922).

‡ 'Roy. Soc. Proc.,' A, vol. 108, p. 592 (1925).

§ 'Roy. Soc. Proc.,' A, vol. 113, p. 420 (1927).

‖ 'Roy. Soc. Proc.,' A, vol. 113, p. 368 (1926); vol. 115, p. 528 (1927), and some unpublished P, R data supplied by Prof. Richardson.

¶ 'Roy. Soc. Proc.,' A, vol. 114, p. 293 (1927).

Table I—(continued).

λ					
4770·54*	6·1	2·8	2·2		—
69·50	2·7	—	—		—
68·12*	3·7	—	—		2α0Q(1)
67·91*	3·5	—	—		—
67·24	4·4	—	—		—
63·82	26·3	9·9	2·7		II BδQ(2)
60·06	2·8	—	—		—
56·92	5·5	2·0	2·8	D ?	—
51·59	4·6	1·5	3·1		—
50·61	1·5	—	—	D ?	4β4Q(3)
45·94	1·6	—	—		4β4Q(2)
43·40	12·4	4·1	3·0		—
42·75	15·4	5·6	2·8		4β4Q(1) II BαQ(2)
42·08	6·0	1·8	3·3		$_{3}\beta_{3}$P(4)
40·98	12·0	4·2	2·9		—
32·81	4·0	1·5	2·7		—
30·77	5·2	1·6	3·2		—
28·90*	3·4	—	—		—
28·51*	2·8	—	—		—
26·13	1·5	—	—		3δ5Q(1)
24·81	9·5	3·4	2·8		—
23·89	5·0	—	—		—
23·01	93·6	22·2	4·2		—
21·56*	7·7	2·3	3·4		—
21·05*	3·8	—	—		4β4R(1)
19·02	91·3	30·0	3·0		—
14·34	1·8	—	—		—
13·93	18·7	6·8	2·6		II BαQ(5)
10·05	12·9	3·5	3·7		$_{2}\beta_{2}$P(5)
09·52	33·0	10·8	2·9		1δ3Q(3)
09·12*	2·6				—
08·47	2·9	—	—		—
05·24	2·8	1·3	2·2	D ?	—
02·55*	5·8	2·0	2·9		—
02·28*	2·4	—	—		—
4695·32	± 1	—	—		3β_{3}P(2) 0δ2Q(5)
92·04	11·3	3·4	3·3		0δ2Q(3)
90·16	34·4	10·0	3·4		0δ2Q(2)
88·43	4·4	—	—		3β3Q(3)
86·79	20·3	7·8	2·6		—
86·12	15·2	4·7	3·2		—
84·64	11·7	4·6	2·5		—
83·79	86	27·4	3·1	$_{1}$A$_{2}$Q(1)	3β3Q(2)
82·35	25·0	8·4	3·0		—
81·38	4·8	—	—		—
80·74*	2·5				
80·43	10·0	3·8	2·6		3β3Q(1)
79·11	18·1	6·1	3·0		—
76·86	2·3	—	—		—
75·35*	9·3	3·4	2·7		—
74·9	5·6	2·1	2·7		—
74·53	9·4	3·7	2·6		II BαQ(7)
73·12	7·3	2·4	3·0		—
71·83	3·7	—	—		—
71·33	28·5	10·4	2·7		—
70·68	9·1	4·0	2·3		—
69·95	4·1	1·5	2·7		—
69·25	5·8	1·9	3·1		—
67·80	7·4	2·4	3·1		—
67·05	8·4	3·0	2·8		—

λ					
4665·58	14·9	5·2	2·9		—
62·79	39·4	15·0	2·6		—
61·39	34·7	11·0	3·1		—
60·34	15·6	6·1	2·6		—
60·00*	3·2	—	—		—
58·79	2·9	—	—		—
57·83	5·1	—	—		3β3R (1)
54·01	9·1	2·7	3·4		—
53·04	37·8	11·6	3·3		—
50·96	2·5	—	—		2β2Q(6)
50·34	3·8	—	—		—
49·47	4·4	—	—		—
48·50	3·8	—	—		—
46·56	2·2	—	—		—
45·33	10·4	3·2	3·4		$_{1}\beta_{1}$P(5)
40·53	2·4	—	—		2β2Q(5)
34·66	—	25·3	—	$_{2}\beta_{2}$P(2)	$_{0}$A$_{0}$Q′(1)
33·95	—	143	—		$_{0}$A$_{0}$Q(1)
31·89	—	70	—		$_{0}$A$_{0}$Q′(2)
31·44	—	30·7	—		$_{0}$A$_{0}$Q(2)
27·96	>100	54·6	—		—
27·05*	3·8	—	—		—
26·79*	2·9	—	—		4γ5Q(2) $_{0}$A$_{0}$Q′(3)
25·61*	8·7	3·0	2·9		—
25·39	32·2	9·4	3·4		$_{0}$A$_{0}$Q(3) 2β2Q(3) 5α2Q(3)?
20·72	5·9	2·3	2·6	D ?	2β2Q(2)
18·32	30·1	10·0	3·0	D	—
17·49	18·8	12·5	1·5	!	$_{0}$A$_{0}$Q′(4) 2β2Q(1)
14·58	3·7	1·1	3·4		$_{2}\beta_{2}$P(1)
13·10	4·0	1·4	2·9		—
07·38	13·7	5·2	2·7		—
05·35	4·4	1·6	2·8		$_{0}$A$_{0}$Q(5)
4598·48	7·6	2·5	3·0		3γ4Q(4)?
97·22	6·4	1·9	3·4		—
95·13	2·4	—	—	D	—
91·84	5·5	1·8	3·0		$_{2}\gamma_{2}$P(3)
89·92	1·7	—	—		3γ4Q(1)
88·67	6·8	2·1	3·2	D	—
85·09	3·8	1·5	2·5		—
84·48	10·8	3·6	3·0		$_{1}\gamma_{2}$P(4)
83·88	4·1	1·6	2·6		$_{0}\beta_{0}$P(5)
82·58	79·5	21·5	3·7	$_{2}\beta_{2}$R(2)	II AαP(4)
81·53	13·8	5·0	2·8		—
80·04	>100	61	—	$_{2}\gamma_{2}$P(2)	II AαP(5
79·43	27·8	8·1	3·4		—
78·00	39·5	8·8	4·5		II AαQ(2)
77·23	4·0	1·4	2·9		1β1Q(5)?
76·54	2·9	1	2·9	D	—
75·86	43·5	11·8	3·7		II AαR(2)
72·73	—	35·1	—		II AαQ(3)
72·19*	—	1·8	—		$_{1}\beta_{1}$P(2)
70·89	2·1	—	—		2β2R(3) 2γ3Q(5)
68·8*	2·5	—	—		1β1Q(4)
68·10	>100	46·1	—		II AαR(3)
65·56	4·2	1·2	3·5		—
64·76	1·2	—	—		2γ3Q(4)
63·68	13·6	3·7	3·7		II AαQ(4)

Table I—(continued).

λ					
4562·24	17·2	4·8	3·6		1β1Q(3)
61·13	4·2	1·2	3·5		—
60·20	2·4	0·9	2·7		$_2\beta_2$R(4)
58·91*	5·3	—	—		
58·62*	14·9	4·8	3·1	$_0\beta_0$P(4)	II AbQ(2)?
58·26*	12·1	3·3	3·7		
57·37*	9·8	3·8	2·6		1β1Q(2)
57·10*	7·9	3·4	2·3		2γ3Q(2)
54·13	41·0	15·4	2·7		1β1Q(1)
52·73*	2·0	—	—		—
52·49	4·1	1·2	3·4		$_1\beta_1$P(1)
51·74*	5·3	—	—		—
51·42*	3·3	—	—		—
50·96	31·0	—	—		II AaQ(5)
49·91	9·7	—	—		—
47·93	6·6	—	—		$_0\gamma_1$P(3)
47·26	3·3	—	—		—
44·12*	1	—	—		—
43·68	13·7	—	—		—
41·12	4·5	1·2	3·7	$_1\gamma_2$P(2)	II AaR(5)
39·15	15·8	6·0	2·6		II AcQ(2)
38·28	10·7	3·3	3·2		—
37·83	6·3	2·2	2·9	D ?	—
35·88	4·9	1·4	3·5		—
34·56	22·3	8·2	2·7		0β0P(3)
34·17	18·3	6·8	2·7		II AcR(3)
33·08	15·4	5·3	2·9		$_1\beta_1$R(1)
31·97	4·6	1·7	2·7		1γ2Q(5)
31·14	4·2	—	—	D ?	—
29·09	11·5	5·0	2·3	D ?	—
27·74	3·3	1·1	3·0		—
27·18	4·3	1·3	3·3		1γ2Q(4)
24·95	1·4	—	—		—
24·13	18·0	6·4	2·8		—
23·18	6·5	2·4	2·7	Tr ?	1γ2Q(3)
22·37	1·6	—	—	$_1\beta_1$R(2)	0β0Q(6)

λ					
4521·44	12·1	3·6	3·4		—
20·64	1·4	—	—		1γ2Q(2)
19·95	5·7	1·7	3·3		II AdQ(2)
19·13	4·5	1·4	3·2		1γ2Q(1)?
17·87	2·8	1	2·8		—
17·47	4·2	1·6	2·6		—
15·56*	5·5	—	—		—
15·26*	3·6	1·3	2·8		II AbQ(6)?
14·34	6·9	1·9	3·6		II AaQ(7)
13·79	3·2	—	—		—
12·81	4·0	1·2	3·3	D	$_0\beta_0$P(2) $_1\beta_1$P(3) 0β0Q(5)
11·70	9·3	3·7	2·5		—
10·91	13·0	4·2	3·1		—
09·13	3·0	1·1	2·7		—
05·63	14·9	4·2	3·5		$_0\gamma_1$P(2)
04·90	3·1	—	—		—
04·55	3·7	1·2	3·1		—
04·07	1·6	±0·5	±3		0β0Q(4)
02·68	4·2	1·6	2·7	D ?	—
01·91	14·3	5·2	2·7		—
00·11	3·3	—	—		—
4499·57	3·2	0·9	3·5		—
98·52	26·1	6·3	4·1		0β0Q(3)
98·10	68·0	25·4	2·7		—
97·6*	8·0	3·5	2·3		—
97·10	5·6	—	—		—
96·69	2·6	—	—		—
93·67	12·2	4·2	2·9		0β0Q(2) II AdQ(5)
90·45	48·0	16·6	2·9		0β0Q(1) II AaQ(8) $_0\beta_0$P(1)
87·80	—	6·6	—		—

Remarks.

There are many lines we have found double which Merton and Barratt give as simple or "probably double," and in such cases we have measured the wave-lengths of each of the components with the help of neighbouring simple lines. These "doublets" are indicated with the sign *.

It seemed possible that not all the lines we have measured might be sharp; in that case the maximum of the line would be too low and therefore we should find a value for the intensity which is too low; the half-width would be too large in those cases. In order to control this phenomenon we have measured the half-width of the lines at all parts of the spectrum, but within the error of experiment these half-widths agreed and we know our plate to be sharp.

When working with a grating the appearance of ghosts can influence the results. As a control we have therefore measured on a special plate of the

Hg spectrum the proportion of the intensities of ghosts and principal line; in this spectrum we have found for the proportion 1 : 150; therefore only for the strongest lines could the ghosts have any influence. However, in no case is any influence present.

For comparison we give in Table II some of the results of McLennan and the corresponding figures obtained in our measurements. Column 1 gives the wave-lengths, column 2 our results, column 3 the results obtained by McLennan and column 4 the proportion between them, which should be approximately constant if the method used in their work was correct. One can see that really the proportion is constant when the intensity of the lines is moderate, but that in all cases where the intensities are only a little greater the proportion is not at all constant.

Table II.

λ				λ			
4767·24	4·4	4½	1·0	4732·81	4·0	3½	1·1
63·82	26·3	11	2·4	30·77	5·2	5	1·0
60·06	2·8	3	0·9	23·01	93·6	26	3·6
56·92	5·5	4½	1·2	4674·53	9·4	16	0·6
51·59	4·6	3½	1·3	73·12	7·3	11	0·7
42·76	15·4	17	0·9	71·33	28·5	23	1·2
40·98	12·0	9½	1·3				

Summary.

With the help of the Utrecht method of intensity measurement, for the region 4500–4900 Å., the intensities of lines in the secondary spectrum of hydrogen are measured and compared with figures given by McLennan and others.

One of us, Dr. Kapuscinski, has been enabled to co-operate in this work by a Rockefeller Fellowship. He takes this occasion to express his gratefulness to the International Education Board.

An Experimental Study of the Flow of Water in Pipes of Rectangular Section.

By S. J. Davies, M.Sc., M.I.Mech.E., and C. M. White, B.Sc., of King's College, London.

(Communicated by E. V. Appleton, F.R.S.—Received February 11, 1928.)

During the course of study, by the authors, of the flow of fluids in the small clearances which exist between the moving and fixed parts of certain machines, an accurate knowledge was desired of the range over which the equations of viscous flow could be applied. An exhaustive search revealed an absence of any record of experimental work which could be of direct assistance. An investigation was accordingly undertaken with the object of obtaining the desired information, and as the preliminary results were of an interesting and unexpected nature, the experiments were extended to cover the whole range of velocities and dimensions permitted by the apparatus. They have shown briefly that the lower critical velocity (as ordinarily understood) for flow between flat plates occurs at a value of the Reynolds number about one-half* that found for pipes of circular cross section, if the linear dimension in that number is the distance between the plates and the diameter respectively. For velocities well below this limit there is evidence, however, of a distinct deviation from true viscous flow if initial disturbing factors are present, and the influence of such disturbing factors does not disappear entirely until a second well-defined limit is reached, which has a value of about one-tenth of the lower critical number. It would appear that below this limit eddies do not exist at any point in the pipe, and the flow is truly viscous. The suggestion is accordingly made that there may be three distinct types of flow: (*a*) one in which eddies cannot exist, corresponding to truly viscous flow; (*b*) one in which eddies may exist, due to an initial disturbance, but cannot be sustained in the pipe, the initial eddies therefore ultimately disappearing; and (*c*) one in which eddies once generated will be maintained without decrement throughout the pipe, corresponding to truly turbulent flow.

---

* This result would follow if the hydraulic mean depth be substituted for the lateral dimension in the Reynolds number. For turbulent flow in smooth pipes this substitution is justified by the experimental results obtained by Schiller ('Z. f. ang. Math. u. Mech.,' vol. 3, p. 2 (1923)) from tests of a very wide range of pipes of various shapes. Reynolds stated that a value of about one-half was to be expected, but he does not appear to have explained the grounds upon which he based the statement.

The use of a channel of rectangular cross section for a study of the fundamental laws of the flow of fluids possesses advantages, in point of simplicity, which were recognised at once by Reynolds in his classical research into the cause of instability of flow. In the form in which this channel is used by the present writers, an additional and important advantage is obtained over the circular pipe by the fact that the controlling dimension may be varied over a wide range whilst retaining the same surfaces as boundaries. It is, in essentials, an adjustable pipe. The upper plate A (fig. 1) and the lower plate B are brass castings suitably drilled to provide inlet and outlet passages and pressure measuring points. The surfaces forming the pipe are hand scraped to a surface plate, and are separated at the ends by brass foil shims of suitable thickness, thus providing a passage between the inlet and outlet ports. The sides of this passage are closed by the plates C and D, very thin rubber insertion providing a watertight joint. All parts are sufficiently robust to reduce distortion under pressure to an amount found to be negligible.

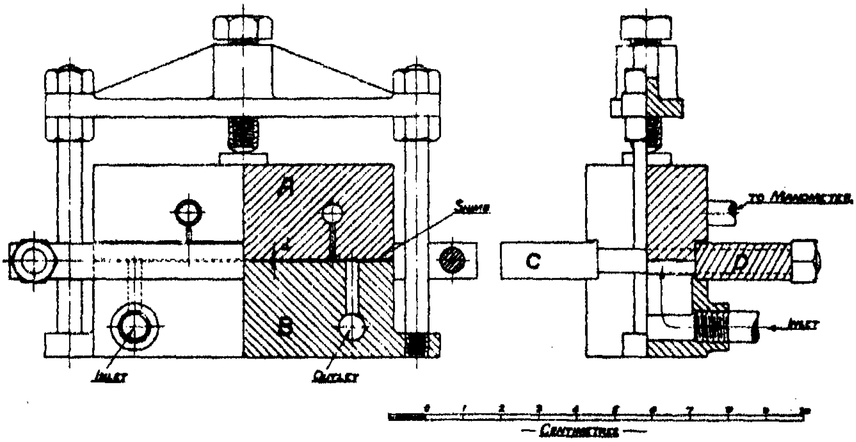


Fig. 1.—" Built-up " Pipe of Rectangular Section.

The water under pressure was taken from the mains, with suitable arrangements to ensure sufficient constancy of flow. The water flowing in a definite time was weighed. Pressure differences were measured by a simple manometer, employing either mercury or water as fluid according to the range of pressure. The temperature of the water was observed by a thermometer placed in the outlet stream near the test pipe. Difficulty was experienced in keeping the temperature constant, and it was thought possible that variations might cause inaccuracy. To check this, the temperature of the water was controlled by means of an electrical heater, and some special tests were made at constant pressure, but with the temperature deliberately varied; that is to say, the flow was varied by controlling the viscosity. The region of the critical velocity was selected for these tests, thus including the transition from viscous to turbulent flow, during which the frictional drag on the walls of the pipe increases more rapidly than the square of the velocity. This region is generally thought to be particularly unstable, and it was hoped that any discrepancies would be at once apparent. It was interesting to observe, with constant pressures, the paradoxical effect of a decreasing velocity resulting from a decreasing viscosity, and far from being unstable, these conditions could be reproduced with certainty. No discrepancies due to temperature changes could be detected.

With regard to the accuracy of measurement, the errors of the water quantities may be taken as less than 0·2 per cent. It is difficult to assess the magnitude of the possible errors of temperature measurement. The thermometer was read to 0·1° C., but the temperature of the water in the test pipe might have been different. It seems unlikely that the error would exceed 0·5° C., which would correspond to a possible, but unlikely, error of 1·5 per cent. in the value of the viscosity. The manometer readings for the most part should not involve errors exceeding 0·2 per cent. (a few readings with large flows through the smallest pipe, and likewise a few with small flows through the largest pipe, may be subject to a greater error). Every care was taken to see that no burrs were left at the gauge points, and that there were no leaks nor air bubbles in the manometer connections.

In any particular series of observations it was found that a test could be repeated within 0·3 per cent. of the previous value.

The following notation is used :—

B = Breadth of section.
d = Distance between surfaces, that is, depth of section. (Where reference is made to pipes of circular section, d denotes diameter.)
F = Intensity of frictional drag on wall of pipe.
g = Acceleration due to gravity.
L = Distance from inlet to pipe to first measuring point.
W = Mass flow in unit time.
μ = Viscosity.
ρ = Density.
v = Mean velocity, *i.e.*, $W/(\text{area} \times \rho)$.

The range of values of the Reynolds number, $\rho vd/\mu$, which it was possible to investigate with the apparatus extended from 60 to 4600. A record of the conditions under which the tests were carried out is given in the table below. Preliminary tests of restricted range, such as those mentioned above in connection with temperature variations, are omitted from the table, as they were intended only to determine the limitations of the apparatus.

Series.	Measured depth.	Calculated depth *d*.	Range of $\frac{\rho v d}{\mu}$.	No. of tests in the series.	Approximate temperature.
	cm.	cm.			° C.
1	$0 \cdot 015_1$	0·0154	64 to 1460	10	16·7
2	$0 \cdot 020_5$	0·0205	230 ,, 1640	20	8·6
3	$0 \cdot 020_8$	0·0212	75 ,, 2200	20	16·0
4	$0 \cdot 023_1$	0·0236	87 ,, 2500	22	14·8
5	$0 \cdot 024_2$	0·0244	210 ,, 1830	28	6·0
6	$0 \cdot 035_0$	0·0344	71 ,, 3240	16	12·7 to 22·6
7	$0 \cdot 036_8$	0·0364	390 ,, 2260	13	8·0
8	$0 \cdot 040_9$	0·0412	60 ,, 3640	26	15·5
9	$0 \cdot 047_7$	0·0476	69 ,, 4180	50	14·6 to 15·6
10	$0 \cdot 068_1$	0·0681	111 ,, 4620	15	15·8 ,, 17·4

Breadth of section, B	2·54 cm.	for series 1 to 10.
Test length	3·81 cm.	
Entrant length, L	1·25 cm.	

The general results of the tests are given in fig. 2, in which log $(F/\rho v^2)$ is shown plotted in the usual manner on a base of log $(\rho v d/\mu)$ but limited to values of $\rho v d/\mu$ greater than 355. It will be seen that, excepting series 7, 8, 9 and 10, all the points fall on a common curve, the general character of which is similar to that obtained from tests of smooth circular pipes.* The straight line, at a downward slope of 45 degrees, which represents viscous flow, is adhered to very closely indeed for values of $\rho v d/\mu$ up to about 890. A small deviation is then noticeable, but it is not until about 1400 is reached that the curve begins to rise. The subsequent rise is somewhat indefinite, and the exact position of the lower critical point cannot be seen. (It is shown later that the lower critical point lies at 1440.) Above 2400 the points of all series lie satisfactorily on a line which is in reasonable agreement with that for smooth circular pipes. With regard to the series 7, 8, 9 and 10, it will be seen that deviation from the viscous line occurs earlier than in the remaining series. This is shown more clearly by a different method of plotting.

The method of plotting in fig. 2, although generally satisfactory as a means of representing the law of resistance, had distinct drawbacks in the present tests, where deviation from theoretical viscous flow is under consideration. The straight line of viscous flow in fig. 2 has a downward slope of 45 degrees, so that, with the scale of plotting practicable, small differences of slope could not be easily observed.

---

* *Cf.* Stanton and Pannell, 'Phil. Trans.,' A, vol. 214, p. 199 (1914).

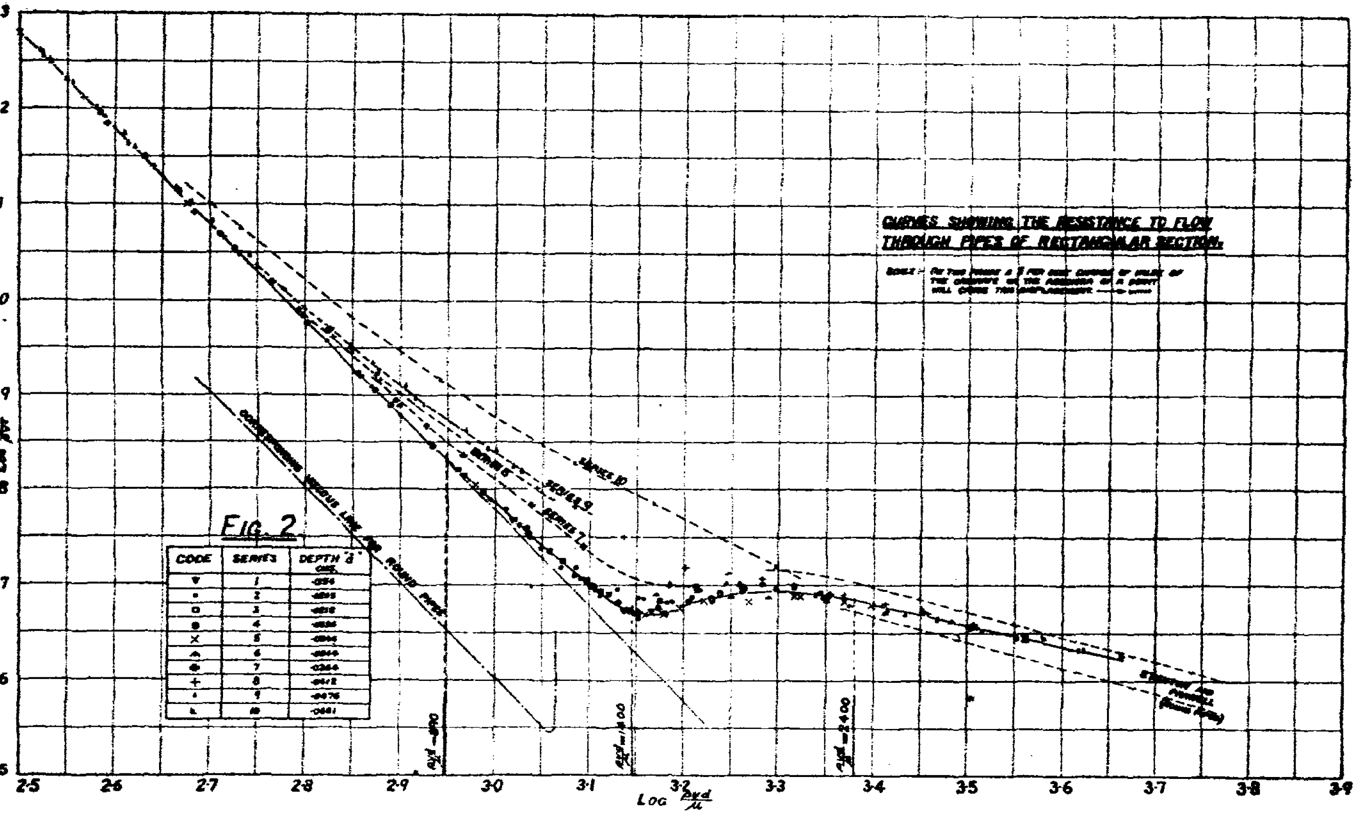
The characteristic of viscous flow is that stresses are in direct proportion to the rate of shear. This would suggest F/v as the basis for the ordinate. Multiplying by d and dividing by μ renders this quantity dimensionless. For purely viscous flow, the resulting function $Fd/\mu v$ is constant, and gives a horizontal straight line when plotted logarithmically on a base of $\log (\rho vd/\mu)$. It may be noted that $Fd/\mu v$ is the product of $F/\rho v^2$ and $\rho vd/\mu$. Fig. 3 is plotted in this way, and with the relatively large vertical scale possible, slight deviations from purely viscous flow are readily detected.

The expressions $F/\rho v^2$ and $Fd/\mu v$ are both functions of d, and before they can be evaluated it is necessary to determine d. The direct measurement of such a small and inaccessible dimension is difficult and liable to error. The difficulty was overcome by the use of a calculated value of d. The theoretical equation, $F/\rho v^2 = 6\mu/\rho vd$,* has been assumed to be true, and d has been calculated from flows that are definitely viscous. The criterion, $\rho vd/\mu$, which reduces to $W/\mu B$, is, for any given value of W, independent of d, an advantage peculiar to flow between parallel plates. Thus in figs. 2 and 3 both ordinates and abscissæ are independent of any directly measured values of d. Table I gives the values of d calculated in this way and those found by direct measurement. In every series the two values lie within the range of accuracy of the direct measurement, so that it is most improbable that any appreciable distortion of the curves could have been introduced by the substitution of the relatively precise calculated value for the necessarily rough value obtained by measurement. The method of working out the results is further explained and a table of typical results is given in the Appendix.

In fig. 3 it will be seen that the three series of tests in which the value of d was least, namely, 1, 2 and 3, the point of deviation, that is, from the horizontal line which represents viscous flow, occurs at the same value of $\rho vd/\mu$, namely, 890. As, however, the value of d is increased, the value of $\rho vd/\mu$, at which deviation first occurs, diminishes progressively. There is, therefore, a departure in these cases from true similarity of flow. The only variable to which it appears possible to ascribe this departure from similarity is the ratio of what may be called the "entrant length," L, to the depth, d. Since the entrant length is constant for all tests, the ratio L/d varies with each change of depth,

---

* This is a more convenient form of the usual equation, $Q = \frac{2}{3} \cdot \frac{h^3}{\mu} \cdot \frac{dp}{dl}$. *Cf.*, for example, Gibson, 'Hydraulics and its Applications,' p. 67 (Constable, 1919).


CURVES SHOWING THE RESISTANCE TO FLOW THROUGH PIPES OF RECTANGULAR SECTION.
FIG. 2.
CODE
SERIES
DEPTH "d" cms.
1
2
3
4
5
6
7
8
9
10
·054
·0412
·0476
SERIES 10
SERIES 8 & 9
SERIES 6
SERIES 7 & 8
CORRESPONDING VISCOUS LINE FOR ROUND PIPES
STANTON AND PANNELL (Round Pipes)
$\frac{\rho v d}{\mu} = 890$
$\frac{\rho v d}{\mu} = 1400$
$\frac{\rho v d}{\mu} = 2400$
$\text{Log}\ \frac{\rho v d}{\mu}$
$\text{Log}\ \frac{R}{\rho v^2}$
2·5
2·6
2·7
2·8
2·9
3·0
3·1
3·2
3·3
3·4
3·5
3·6
3·7
3·8
3·9
$\bar{2}$·3
$\bar{2}$·2
$\bar{2}$·1
$\bar{2}$·0
$\bar{3}$·9
$\bar{3}$·8
$\bar{3}$·7
$\bar{3}$·6
$\bar{3}$·5

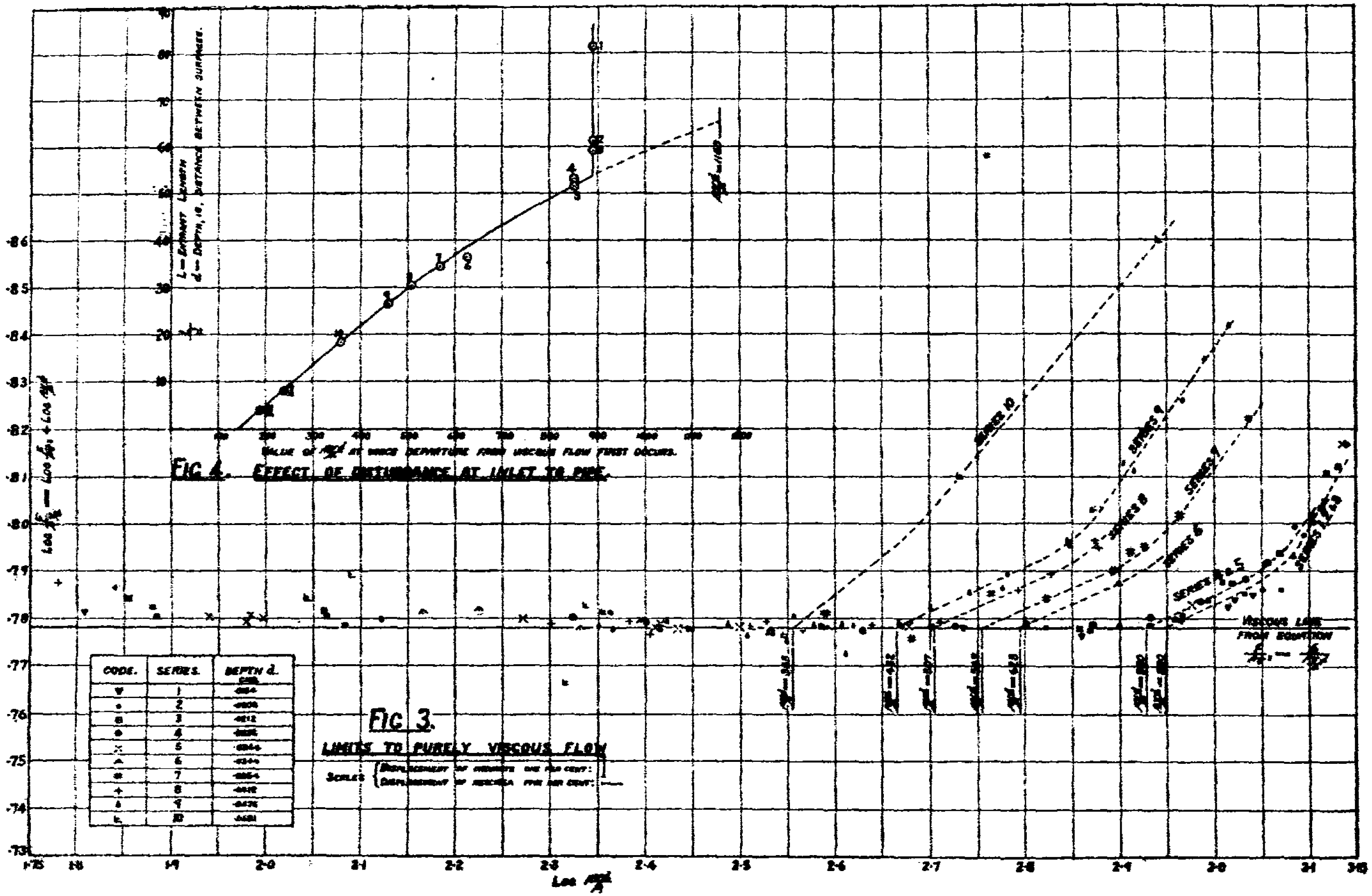


FIG. 4. EFFECT OF DISTURBANCE AT INLET TO PIPE.

FIG. 3. LIMITS TO PURELY VISCOUS FLOW.

and therefore true geometrical similarity is not secured. When, however, as in fig. 4, the value of L/d is plotted against the value of $\rho vd/\mu$, at which deviation from the viscous line first occurs, two very interesting facts emerge. In this figure the vertical portion of the curve, at a value of $\rho vd/\mu = 890$, represents the conditions in those series in which d was less than 0·0236 cm. The remainder of the points lie upon a curve which, if continued for a short distance, would appear to cut the horizontal axis at a value for $\rho vd/\mu$ of 140. The curve meets the vertical portion in a sharp discontinuity when the value of L/d is 54. It is therefore evident that when L/d is greater than 54, it ceases to have any effect upon the position of the first deviation from true viscous flow.

The lower part of the curve appears to indicate, for each pipe, the necessary values of $\rho vd/\mu$ for which the eddies, generated by discontinuities at the entrance, are caused to be transmitted along the pipe through a distance just equal to the entrant length, before true viscous flow is attained. In short, it gives the distance, in terms of d, through which the eddies will be transmitted before they disappear completely. Schiller* found, for circular pipes, that an entrant length of 130 diameters was necessary to give freedom from inlet disturbances at a value of $\rho vd/\mu$ of 2320.

The positive value of $\rho vd/\mu$, at which a backward extension of the curve cuts the axis of zero L/d, suggests that, even with zero entrant length, truly viscous conditions will obtain over a certain range of velocity. In other words, it indicates that whatever may be the pre-entrant conditions, eddies cannot be transmitted to, or exist in, the channel when the Reynolds number is less than the value given by the intersection, namely, 140. This may well be a critical value in connection with viscous flow.

In view of the importance of such a conclusion, it was felt that confirmation of the form of the lower part of the curve was desirable. The departure from viscous flow manifests itself primarily as a difference of total kinetic energy of flow at the two gauge points. At the very small values of $\rho vd/\mu$ at which confirmation was desired, this kinetic energy is extremely small in comparison with the frictional losses, and the point of departure is consequently difficult to locate. The difficulty was overcome with the aid of an additional gauge point, which reduced the measuring length to 0·7 cm. The entrant length was reduced by reversing the direction of flow through the pipe, thus altering, as far as possible, the conditions of the previous tests.

Two series of tests were made, with entrant lengths of 0·2 and 0·1 cm. respectively, and both with a calculated depth d of 0·0238 cm. With these conditions the points I and II (fig. 4) were obtained. It should be noted that, while there is little doubt about the exact value of point II, in the other series the departure was not so clearly defined, and point I may be plus or minus 10 per cent. from the value given. It will be seen that these tests confirm, in a remarkable manner, the general form of the curve in fig. 4, and show conclusively that the apparent lack of similarity of flow in the earlier tests is due entirely to the variation in the L/*d* ratio.

---

* 'Z. f. ang. Math.,' vol. 1, p. 436 (1921).

Evidence of the existence of a critical point below the ordinary lower critical velocity is to be found in the published results of other investigators. Sorkau,* in the course of the determination of viscosities of various fluids, found discontinuity in the curve of discharge for capillary tubes in the region of $\rho vd/\mu = 430$. Carothers,† from large-scale tests with oil flowing through jointed pipes, found a critical point of the order of 200. Grindley and Gibson‡ also found a critical point of this order, in tests with air flowing through a long coiled lead pipe. In each of these cases it was observed that the resistance varied according to a power of the velocity which did not differ materially from 1·2 for flow above the respective critical points. Thus, it is evident that the change in the nature of the flow was distinct from the change that is customary at the ordinary lower critical point. There is, therefore, support for the view that the critical point suggested by the present tests is general in character, and occurs with flow through pipes of circular section.

Further, the lower critical point, as generally understood, is definite in character and in value, and is not even affected by considerable roughness of the pipe.§ For circular pipes its value may be accepted as 2300 very closely. Analytical treatment, however, gives values for instability in circular pipes at so low a figure as 180.|| The recognition of a third critical point might explain this disagreement between theory and experiment. The lower critical point is perhaps not associated so much with the stability of an eddy, as such, as with the self-maintenance or persistence of a series of eddies throughout the length of the pipe. An eddy may be produced at the entrance to the pipe, or at some obstacle within the pipe, but it may never reach the gauge points. In the analytical treatment, it is the possibility of the formation of the eddy, rather than its subsequent history, that receives attention. Thus the criterion of stability, as found analytically, has little in common with the lower critical

---

* 'Phys. Z.,' vol. 14, p. 759 (1913).

† 'Roy. Soc. Proc.,' A, vol. 87, p. 154 (1912).

‡ 'Roy. Soc. Proc.,' A, vol. 80, p. 114 (1908).

§ Schiller, 'Z. f. Physik,' vol. 3, p. 412 (1920).

|| Orr, 'Proc. R. Irish Acad.,' vol. 27, A, Nos. 2, 3 (1907).

velocity, as observed experimentally. The first departure from viscous flow in a very rough pipe, or near the abrupt entrance to a smooth pipe, or even in a curved pipe, may correspond to the theoretical case, whereas the persistence of a stream of eddies is clearly a different phenomenon. The agreement between the values from the experiments quoted above and those obtained theoretically by Orr (namely, 180 for circular pipes and 117 for flow between flat plates) is sufficiently close to lend support to this argument.

Fig. 2 shows a series of distinct curves in the region of transition from viscous to turbulent flow. There can be little doubt that the primary cause is disturbance at inlet to the pipe, but the manner in which such disturbance increases the resistance in the test length is not altogether clear. The action is probably two-fold since: (*a*) a fictitious loss of head is obtained, owing to dissimilarity of the velocity distributions at the two gauge points, and (*b*) the frictional resistance is increased owing to the presence of a certain degree of turbulence. The former effect will depend upon the ratio of the gauge length to the depth d, as well as upon the ratio L/d, whereas the latter effect will, other conditions being the same, be relatively unaltered by change of gauge length. Thus the fact that several curves are obtained is not the result of a breakdown of the general principle of similarity, but is due to the relatively different spacing of the gauge points in the different series.

The dominant factor in the design of the apparatus was the necessity of keeping the overall length of the pipe as short as possible in order to minimise the amount of distortion of the surfaces, caused by the high fluid pressures to be expected with small depths d of the pipe—a distortion which, for a given pressure, varies as the fourth power of the length. The length between pressure-measuring points should, on the other hand, be as great as possible, in order to give reasonable facility of measurement at the lowest velocities. Thus the entrant length and the exit length were relatively reduced, and in certain series, namely, 6, 7, 8, 9 and 10, the entrant length is insufficient for flows near the lower critical velocity. In the case of the remaining series, however, the entrant length is sufficient, and the slight departure from viscous flow, noticeable in fig. 2, which occurs for flows above 890, must be caused by some other disturbing factor within the pipe itself. The departure increases with increasing flow, and at 1440 the curve begins to rise. The rise is gradual, and it is not until 1750 that transfer to the line of turbulent flow is complete. Tests of circular pipes show a sudden jump to turbulent flow if sufficient entrant length be provided, but if the entrant length be short, then, according to Schiller,* a rounding of the transition takes place. Such a rounding occurs in the present tests, in which, however, it appears to be due to some other cause, since check tests with an entrant length of 140 d gave the same rounded form of curve. Some rounding might be caused by a possible variation of velocity across the width of the pipe, particularly as the resistance to turbulent flow is not very much greater than the theoretical resistance to viscous flow. For circular pipes, the viscous resistance is relatively 25 per cent. less, necessitating a jump at least two and a half times as great as that for a rectangular pipe. The different shape of the curve in the two cases makes comparison of the numerical values for the lower critical point difficult. The value of 2300 for circular pipes relates to the almost vertical rise between the viscous and the turbulent lines. In the present tests the rise begins at 1440, which is definitely greater than one-half of 2300. Thus, it appears that the lower critical point does not occur at a precisely constant value of $\rho vm/\mu$ (where m is the hydraulic mean depth) for pipes of different forms. Its position, then, is not fixed by the law governing turbulent flow, which, indeed, appears to be without direct influence upon the critical point.

---

* 'Z. f. ang. Math.,' vol. 1, p. 436 (1921).

Above 2400 the test points of all series fall on a single curve. For comparison, the results obtained by Stanton and Pannell, from tests of circular pipes, are indicated in fig. 2 by dotted lines, which mark the limits within which their points lie. The actual values of $\rho vd/\mu$ for the circular pipes have been divided by two in order to make comparison on the basis of hydraulic mean depth. It will be seen that agreement in value is good. The widest rectangular channels that have been tested previously are those of Fromm,* with a maximum width-depth ratio of 20, which are also in good agreement with the results from circular pipes. The present tests, with a width-depth ratio up to 100, therefore extend the range of proof of the validity of the hydraulic mean depth as a basis of comparison.

It has already been explained that a calculated value of d was used in obtaining values of $F/\rho v^2$. The agreement between the results for the different series, therefore, does not necessarily show that the pipes are smooth, nor that they are equally rough, but it does definitely show that the effect of the character of the walls upon turbulent flow was identical with its effect upon viscous flow. That this effect is too small to be detected in the present experiments, in spite of the large range of ratio of comparative roughness of 21 to 68 (series 1 and 2 do not come into consideration here), is difficult to explain unless the surfaces may be regarded as smooth hydraulically. The idea of similarity is based on the conception of geometrical similarity of motion. Thus, if the surfaces be rough enough to influence the motion, an alteration in the ratio of the size of the irregularities to the distance between the surfaces must constitute a change of roughness. On the other hand, it is difficult to regard the surfaces as smooth, since irregularities of the surfaces, of the order of 2 per cent. of the distance between the surfaces, would, if present in a larger pipe, be regarded as forming an appreciably rough boundary, and the irregularities present were probably of this order. This argument would lead to the conclusion that what may be termed hydraulic roughness does not depend primarily upon the physical roughness of the surface. The experiments of Froude* and others upon the effect of the nature of the surface upon resistance are, however, in direct conflict with this conclusion ; so, clearly, there must be some limit to the size, and possibly to the form, of the irregularities to which it can apply. Presumably such a limit could be specified in terms of the thickness of the viscous film. It is interesting to note that direct velocity measurements, together with experiments upon heat transfer, suggest that near the boundary of a pipe the motion is of a viscous character, but the experimental evidence available does not conclusively prove the existence of a real viscous film of definite thickness. If, however, such a film does exist, then there are grounds for assuming that its thickness is, for the present tests, of the order of $0 \cdot 04\, d$.† That this exceeds the size of the irregularities may explain their lack of influence upon the flow. It is conceivable that the film adapts itself to the irregularities of the boundary, and presents a flat surface to the turbulent core, but if the irregularities exceed the average thickness of the film, they then begin to influence the flow within the turbulent core. If the length of an irregularity is great compared with the film thickness, then the film is displaced bodily inwards, and so causes change of mean velocity of flow of the turbulent core. Thus, two of the independent variables included in the term "roughness" may be defined, but further experimental evidence must first be sought.

---

* 'Z. f. ang. Math.,' vol. 3, p. 339 (1923).

It will be seen that the slope of the turbulent lines in fig. 2 is less than one in four. The latter slope characterises the test results for smooth pipes of all forms of cross section, over a very wide range of velocities. The present tests, particularly those with small values of d, cover only a limited range of turbulent flow, and it is unwise to draw any definite conclusion. Schiller has called attention to a possible source of error in tests with a high-pressure gradient. Should the water be saturated with air, then the reduction in pressure during passage down the pipe will cause some of the air to be liberated, with a resulting increase of resistance. Unfortunately, at the time the tests were made the writers were not aware of this, and it is possible that the smallness of the slope may be caused partly by liberation of air and partly by boundary roughness just beginning to influence the flow. Both these effects, however, will decrease with increase of d. Therefore, in order to account for the uniformity of the slope for the various pipes, *i.e.*, with variation of d, there must be some additional cause which increases with increase of d. This may be found in the relative decrease of the entrant length with the larger values of d. Thus, there are several possible sources of error, each of which, within the range of the tests, appears to be small; and, though the combined effect is not sufficient to alter the value of the resistance coefficient appreciably, yet their systematic nature permits such errors to be accepted as the cause of a noticeable difference of slope.

---

* 'British Association Report' (1872).

† Gibson, in 'The Mechanical Properties of Fluids' (Blackie, 1923), summarises experimental data concerning the resistance to heat transfer. To explain the observed resistance, it is necessary to assume a value between $0 \cdot 38$ and $0 \cdot 14$ for the ratio of the velocity of the inner surface of the film to the mean velocity of flow. It is probable that this ratio increases with decrease of $\rho v d/\mu$. Based on the higher value, a film thickness of $0 \cdot 04\, d$ is not unreasonable for $\rho v d/\mu = 3000$.

Summary.

In all, 400 tests have been made to determine the law of resistance for pipes, varying in section from 2·54 cm. broad by 0·0154 cm. deep to 2·54 cm. broad by 0·0681 cm. deep. A range of $\rho vd/\mu$ from 60 to 4600 has been investigated.

By employing progressively shorter "entrant lengths" evidence has been found suggesting the existence of a third or "lowest" critical point, in the neighbourhood of $\rho vd/\mu = 140$, below which eddies are not transmitted along a pipe.

A curve showing the approximate distance necessary for eddies to die out has been obtained for a range of $\rho vd/\mu$ from 140 to 900.

For turbulent flow, it has been found that the values of the resistance coefficient are the same as those obtained from tests of smooth circular pipes, provided that the comparison be made on the basis of the hydraulic mean depth. The validity of the latter as the controlling dimension is thus extended to include channels with a width-breadth ratio as great as 100 to 1, and to values of the hydraulic mean depth as small as 0·01 cm. (equivalent to a round pipe 0·04 cm. in diameter).

It has been confirmed that roughness, constituted by irregularities some 2 per cent. of the distance between the surfaces, has no measurable effect upon

the resistance offered to viscous flow, nor could the influence of this roughness upon the value of the turbulent resistance be detected.

The tests were carried out in the Engineering Laboratories at King's College, London. The authors would like to express their best thanks to the Dean of Engineering, Prof. Gilbert Cook, D.Sc., for the facilities placed at their disposal, for his kindly interest in the work, and for many helpful suggestions.

APPENDIX.

For viscous flow between infinitely wide parallel flat plates, the law of resistance is

$$F/\rho v^2 = 6\mu/\rho v d\,; \tag{1}$$

this may be written

$$Fd/\mu v = 6. \tag{1A}$$

Now,

$$F = \frac{dp}{dl}\cdot\frac{d}{2} = \rho g\cdot\frac{h}{l}\cdot\frac{d}{2}, \quad \text{and} \quad v = \frac{W}{\rho B d}.$$

Thus, eliminating F and V from (1A),

$$d^3 = 12\mu W l/B\rho^2 g h \tag{2}$$

during any one series of tests: the only terms in this equation which vary from test to test are:—

W, the mass flow,

h, the difference of head between the gauge points, and

μ/ρ, the kinematic viscosity.

The product $\rho h/\mu W$ is constant for viscous flow, but increases rapidly for turbulent flow, thus giving a convenient method of distinguishing between the two.

A preliminary curve of $\rho h/\mu W$ was plotted on a base of $W/\mu B$ (*i.e.*, $\rho v d/\mu$), and, from inspection of the curve, a value of $\rho h/\mu W$ was selected which was typical of viscous flow. This value, substituted in equation (2), enables the calculated value of d to be determined.

The resistance coefficient $F/\rho v^2$ is taken as $Bd^3\rho^2 gh/2W^2 l$.

The above reasoning ignores the side boundaries, the surface area of which for series 10 amounts to 2½ per cent. of the total surface. There is, however, little objection since the small error so introduced into the value of d is at least partially compensated by the subsequent neglect of the effect of the sides in the expression for $F/\rho v^2$.

The dimension d in $\rho v d/\mu$ is eliminated by the substitution of $W/\rho B d$ for v, thus giving the simple expression $W/\mu B$ for the Reynolds number.

I.	II.	III.	IV.	V.	VI.	VII.	VIII.	IX.	X.	XI.	XII.	XIII.	XIV.
Test No.	Duration of test. secs.	Temperature. °C.	Viscosity.* c.g.s. units.	Mass flow. W gm./sec.	Head. h cm. of water.	$\frac{\rho v d}{\mu} = \frac{W}{\mu B}$.	$\frac{\rho h}{\mu W}$. c.g.s. units.	$\frac{Fd}{\mu v}$.	$\frac{F}{\rho v^2}$.	log $\frac{\rho v d}{\mu}$.	log $\frac{\rho h}{\mu W}$.	log $\frac{Fd}{\mu v}$.	log $\frac{F}{\rho v^2}$.
1	349·6	18·3	0·01043	2·596	7·62	98·02	281·5	5·978	0·06099	1·9913	2·4494	0·7766	$\bar{2}$·7853
11	137·7	18·2	0·01046	3·294	9·75	124·0	283·1	6·013	0·04850	2·0934	2·4519	0·7791	$\bar{2}$·6857
10	131·3	18·25	0·01045	3·454	10·18	130·2	282·1	5·992	0·04602	2·1147	2·4504	0·7776	$\bar{2}$·6629
2	255·3	18·3	0·01043	3·554	10·42	134·2	281·3	5·974	0·04452	2·1277	2·4491	0·7763	$\bar{2}$·6486
9	119·1	18·25	0·01045	3·809	11·21	143·5	281·8	5·985	0·04169	2·1571	2·4499	0·7771	$\bar{2}$·6200
12	113·6	18·2	0·01046	3·993	11·80	150·3	282·4	5·999	0·03991	2·1770	2·4509	0·7781	$\bar{2}$·6011
13	129·1	18·2	0·01046	4·215	12·47	158·7	282·8	6·006	0·03785	2·2005	2·4514	0·7786	$\bar{2}$·5781
3	209·2	18·3	0·01043	4·337	12·77	163·8	282·3	5·998	0·03663	2·2142	2·4508	0·7780	$\bar{2}$·5638
4	201·1	18·3	0·01043	4·511	13·29	170·3	282·5	6·001	0·03523	2·2313	2·4510	0·7782	$\bar{2}$·5469
14	118·5	18·2	0·01046	4·591	13·56	172·8	282·5	6·001	0·03472	2·2376	2·4510	0·7782	$\bar{2}$·5406
5	186·5	18·3	0·01043	4·864	14·31	183·7	282·1	5·995	0·03263	2·2640	2·4505	0·7777	$\bar{2}$·5137
6	170·3	18·3	0·01043	5·236	15·46	197·7	283·2	6·017	0·03044	2·2960	2·4522	0·7794	$\bar{2}$·4834
7	155·3	18·3	0·01043	5·842	17·34	220·6	284·6	6·046	0·02741	2·3436	2·4543	0·7815	$\bar{2}$·4379
8	109·2	18·3	0·01043	8·308	24·97	313·7	288·1	6·120	0·01951	2·4965	2·4596	0·7868	$\bar{2}$·2903

* From tables based on mean values collected by Jakob and Erk, 'V.D.I. Forschungsheft,' 267, p. 9 (1924).

A typical set of observations, together with the necessary reduction to dimensionless form, is given below in order to demonstrate the method adopted.

Series II (check tests to confirm form of curve in fig. 4).

l, gauge length	$0\cdot78_7$ cm.
L, entrant length	$0\cdot1$ cm.
B, breadth of pipe	$2\cdot54$ cm.
d, depth of pipe	$0\cdot025$ cm. approx.
Calculated depth of d from column VIII	$0\cdot0238$ cm.
Ratio L/d	$4\cdot2$.

Preliminary plotting shows that $\rho h/\mu\mathrm{W} = 282\cdot5$ sec.^{2}/gm. cm. for viscous flow. Using this value, equation (2) gives $d = 0\cdot02376$ cm., which value is used calculating columns IX and X.

Throughout the calculations the density has been taken as 1 gm./c.c.

On Quantitative Measurements in Frictional Electricity.

By Wallace A. Macky, M.Sc., Auckland University College, New Zealand.

(Communicated by R. Whiddington, F.R.S.—Received February 27, 1928.)

The first record of any quantitative work on frictional electricity is by Peclet[*] in 1834. He investigated the magnitude of charge produced on varying the pressure and velocity of rubbing, and showed that after a certain amount of friction the surface acquired a constant surface density of charge which was independent of the pressure or velocity of rubbing.

These results have been re-affirmed of recent years by Owen, Morris Jones, French, Cardani, Prati, and various others.[†] Those workers have all used a small specimen, metal or insulator, rubbed by a rotating wheel, in some cases 3 feet in diameter.

Two hypotheses as to the generation of these charges have been put forward, one by Helmholtz, the other by Coehn.

---

[*] 'Ann. Chim. Phys.,' vol. 57, p. 337 (1834).

[†] Owen, 'Phil. Mag.,' vol. 17, p. 457 (1909); Morris Jones, 'Phil. Mag.,' vol. 29, p. 261 (1915); French, 'Phys. Rev.,' vol. 9, p. 151 (1917); Cardani, 'N. Cimento,' vol. 123, p. 199 (1922); Prati, 'Atti Accad. Lincei,' vol. 31 (2), p. 478 (1922).

In 1879 Helmholtz[*] suggested that the fundamental cause of frictional electricity is the contact potential difference which exists between any two substances. On this theory the rubbing of the surfaces together merely serves to bring the two surfaces into entire contact, and the maximum charge is obtained when the entire surfaces have been brought into contact with one another.

As the results of experiments with dielectrics rubbed against different liquids, Coehn[†] enunciated his "law," which states that the insulators are arranged in the frictional electric series in the order of their dielectric constants, the substances with the higher dielectric constants being nearer the positive end of the series. In a later publication[‡] this has been widened to state that the quantity of charge produced is proportional to the difference in the dielectric constants of the two rubbers.

Richards[§] published in 1922 results obtained by bringing together special surfaces flat to within ½ wave-length of sodium light. From his experiments with steel and various dielectrics, he states that Helmholtz's theory as to the nature of frictional electricity is upheld, and also Coehn's law if the value 3·1 be assigned to the dielectric constant of steel.

The original intention in the present investigation was to test Richards' results, using pure substances as specimens, and working with the specimens in a vacuum. The preparation of surfaces as flat as those used by Richards is very laborious, and, mainly owing to the necessity of doing this on the spot, only a portion of the original programme has as yet been concluded. The present paper deals mainly with the effect of alteration of the surrounding medium.

Apparatus.

The type of apparatus in which the specimen is rubbed by a rotating wheel was considered too indefinite, and it was decided to work with specimens of approximately the same size, and to do the rubbing by rotating one specimen about its axis perpendicular to the contact face.

With this object, the apparatus shown in the figure with top removed was set up.

A brass box was cast and the various parts fitted inside. The upper specimen (the insulator) is mounted on a double joint, allowing it to lie flat against the surface of the lower specimen. A length of insulating rod connects the double joint to a ground-in taper fitting, which allows the specimen, when *in vacuo*, to be rotated from the outside.

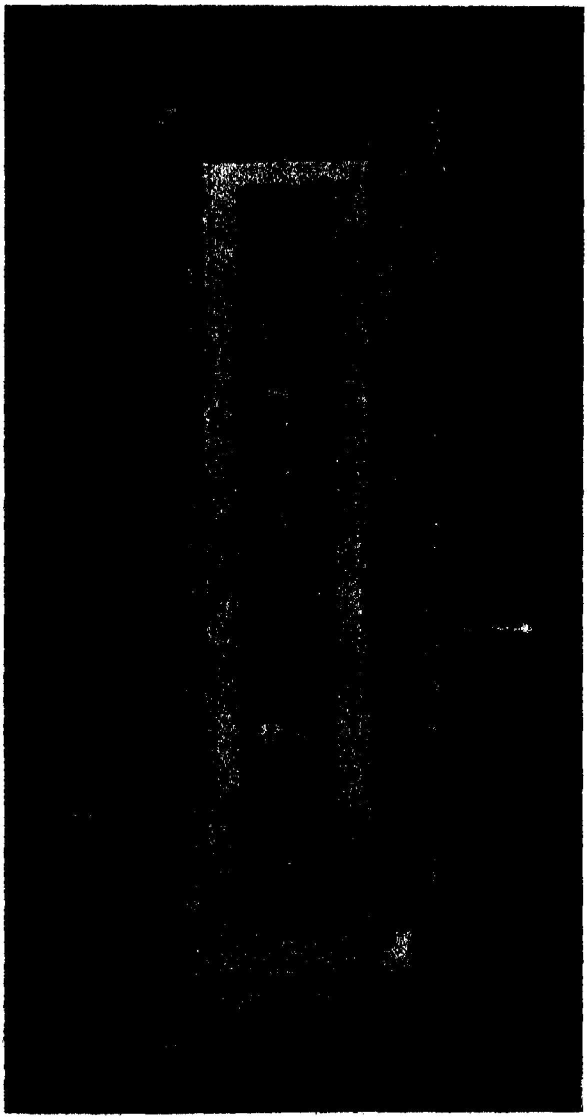
---

[*] 'Wissensch. Abh.,' vol. 1, p. 860 (1894).

[†] 'Wied. Ann.,' vol. 64, p. 217 (1898); and vol. 66, p. 1191 (1898).

[‡] 'Ann. Physik,' vol. 30, p. 777 (1909).

[§] 'Phys. Rev.,' vol. 22, p. 122 (1922).



The lower specimen, metal, is mounted on a rod of insulator screwed to a rack capable of backwards and forwards motion along one direction only. The rack is actuated by a pinion turned through another ground-in taper fitting. This specimen can thus be moved up to make contact with the upper specimen, and then separated to measure the charge.

In the earliest experiments a cylinder mounted on an insulating support was used, connected to an electroscope, and the charge measured by racking down the lower specimen until it was inside the cylinder.

In most of the work, however, the metal specimen has been connected directly to the electroscope by a spring, as shown in the figure. A large ground-in taper fitting in the cover allowed the specimens to be changed without removing the whole top cover. A heating coil was mounted inside the box behind the specimens, and a thermocouple was used to measure the temperature. It was possible to remove all the moving parts from the box through the various openings. To allow the passage of X-rays, a large number of 3-mm. holes were bored in the back of the box behind the heating coil, and a piece of thin copper foil soldered over the outside.

Owing to the large number of soldered joints, it was very difficult to make the box air-tight, so it was surrounded by a wooden case and the space in between filled with paraffin wax. With the oil pump available, a pressure of 1–2 mm. was then obtainable.

Preparation of the Specimens.

The glass specimens used have all been cut from a sheet of plate glass, and in each case have been flat to within 1 or 2 wave-lengths of light, as shown by testing against a standard plate.

Considerable time was spent on the earliest metal specimens, which were ground and polished until flat to within 4–5 wave-lengths of light. Owing to the limited means available, it has not been possible to obtain surfaces optically flat, and experiments have shown that the specially prepared surfaces do not give more consistent results than others polished without taking great care to obtain very flat surfaces. In consequence of this, in all the later experiments the metal specimens have been ground on the finest emery papers, and then polished with rouge on chamois stretched on plate glass. The sulphur specimens have been cast and then ground on fine glass and emery papers stretched on plate glass.

Before each experiment the specimens have been washed in pure ether and allowed to dry.

Methods of Experiment.

1. In the earlier experiments the metal specimen was insulated, and the charge measured by measuring the charge induced when this specimen was lowered into the measuring cylinder.

A gold-leaf electroscope, calibrated by a potential divider with potentials up to 850 volts from a battery of small secondary cells, was used to measure the potential, and the capacity of the system was measured by comparison with a standard condenser.

2. Using the above method, it was not possible to determine the charge on the specimens during the actual rubbing. The measuring cylinder was removed, and the metal specimen connected directly to the electroscope.

This method has shown that all the charge generated is held on the rubbing surfaces, since no deflection of the electroscope is observed until the specimens are separated, when there is a sudden movement which increased to a maximum as the specimens are separated by 2–3 cm.

3. In methods 1 and 2 the specimens were completely discharged with X-rays or radium between each reading. The object in all the later experiments was to obtain the maximum charge possible under the particular conditions, and hence it was necessary to give sufficient rubs to arrive at the flat portions of the curves, as in graph I. In this method, therefore, the specimens were not discharged between each reading, but the metal specimen was isolated at its previous potential before being placed in contact again. It was found that as soon as the specimens were put in contact again the electroscope deflection decreased to zero, and remained so until they were separated. By this method the cumulative effect of all the previous rubs was obtained.

Usually about 10 rubs of 20 revolutions each were given, and the mean of the 10 values taken. The maximum variation from the mean in the results recorded here was from 8 to 10 per cent.

Preliminary Experiments.

Experiments were carried out with method I, and using glass specimens rubbed against metals flat to within 5 wave-lengths of light to test the effect of pressure between the specimens, velocity of rubbing and the amount of rubbing. The results in each case confirmed the work of Peclet and the other workers, namely, that normal pressure and velocity of rubbing have no effect on the maximum charge, but that with small amounts of rubbing a small increase in the amount of rubbing causes a large increase in the charge, which, however, rapidly reaches a maximum value when continued rubbing gives no further increase.

The important point in the results is that in no case did mere contact without rubbing give a charge greater than 0·3 per cent. of the maximum.

The results obtained with four metals are shown on graph I. The charges shown are those on the metals which were negative in each case.

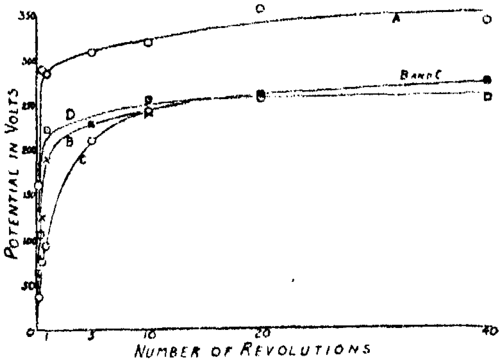


Fig. 1.—Variation in Charge with Number of Revolutions. A, Steel. B, Gold. C, Aluminium. D, Silver.

Variation in the Charge with Decrease of the Pressure of the Surrounding Air.

The apparatus was constructed with the intention of working with the surrounding gas at as low a pressure as possible. It was found, however, that at a pressure of about 2 mm. practically no charge could be obtained, although a large charge was generated when the air was let in. This led to further experiments, which have shown that the magnitude of charge obtained with any two specimens depends very largely on the pressure of the surrounding air.

This result is shown by graphs 2–6,* which show the charges obtained at different pressures with various specimens. The readings were taken at the different pressures in the order shown without discharging the specimens in between.

The graphs show how the charges are decreased after the specimens have been subjected to the low pressures. There was usually an interval of about 10 minutes between the readings at one pressure and those at the next, the whole experiment taking 2½–3 hours. It was found that if the specimens were left undisturbed at the end of the experiment, after a varying period the earlier maximum charge could be obtained, showing a "recovery" of the surface almost certainly due to the adsorption of a gas film. This effect was much more marked in some cases than in others. It is shown clearly in graph 4, where the

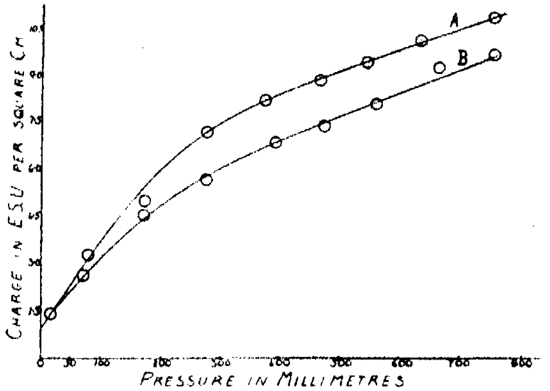


Fig. 2.—Aluminium-Sulphur. A, Decreasing pressure. B, Increasing pressure.

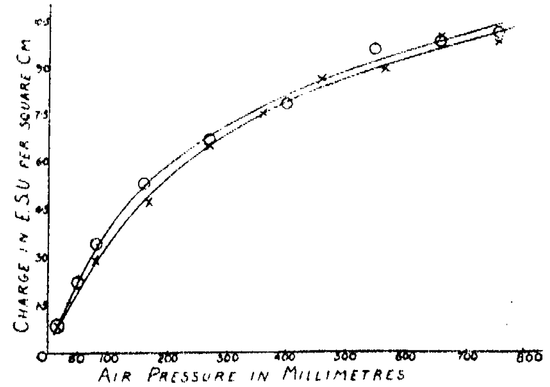


Fig. 3.—Aluminium-Sulphur. ⊙, Decreasing pressure. ×, Increasing pressure.

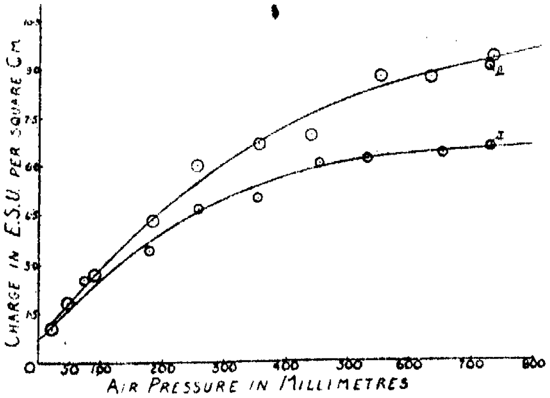


Fig. 4.—Aluminium-Sulphur. ⊙, Decreasing pressure. ⊙, Increasing pressure.
Reading II 30 minutes after Reading I.

recovery took place in half an hour, a shorter time than usual. Graph 5 shows the values at low pressures only, and indicates that the charge is directly proportional to the pressure within these limits. It was found that glass gave less consistent results than sulphur. The values at any pressure on a particular occasion were as constant as with sulphur, but sometimes the values at different pressures did not vary as regularly as was the case with sulphur.

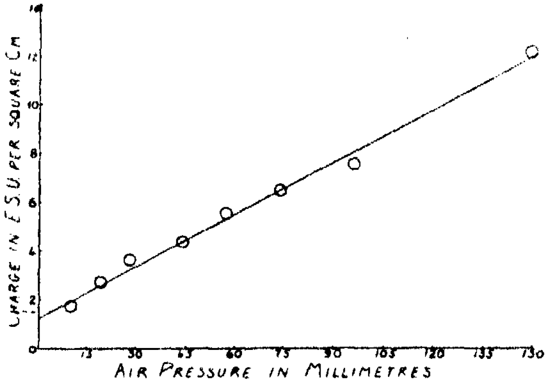


Fig. 5.—Aluminium-Glass.

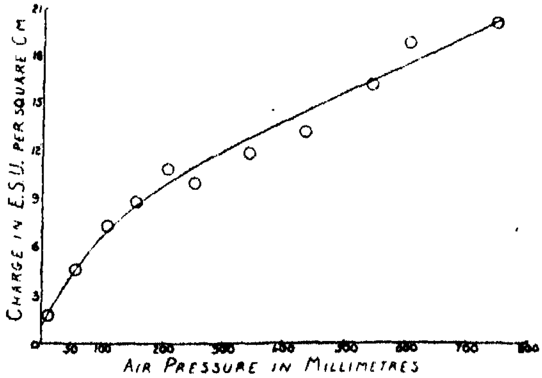


Fig. 6.—Steel-Glass.

The diminution of charge with pressure of the air occurs to an exactly similar extent when gold or silver are substituted for aluminium and steel, and also when the glass is replaced by ebonite. When silk or chamois leather stretched on ebonite are rubbed against steel or aluminium, a much smaller charge is obtained at ordinary pressures than when glass or sulphur is used. The diminution with decreased pressure in these cases is therefore not so marked as with the solid rubbers. It is possible also that the soft materials retain more air at the lower pressures than do hard substances.

The above results show that in every case the magnitude of charge produced is largely dependent upon the pressure of the surrounding air. This effect has been obtained in every experiment carried out, and many more have been done than are recorded above. The important point in these results is that if a charge is generated by rubbing the specimens in air, then, on reducing the pressure and leaving the specimens in contact, this charge is greatly reduced, and becomes almost zero when a pressure of 2–3 mm. is reached.

These experiments were all carried out with dry air surrounding the specimens, and it was necessary to determine if the effect was common to other gases.

The box was filled in turn with sulphur dioxide, carbon dioxide and hydrogen, each carefully dried.

The charge was in each case measured at atmospheric pressure and then at a low pressure. The results obtained are almost identical with those obtained in air, showing that it is some property common to each gas which determines the magnitude of the charge.

Other Experiments on the Decrease in Charge at Low Pressures.

(*a*) In the earlier experiments, using a measuring cylinder and the metal specimen isolated, it was found that if the metal was charged by rubbing in air, then lowered into the measuring cylinder and the box evacuated, the charge on the specimen passed to the cylinder, showing the potential was sufficiently high to cause discharge over a distance of 2–3 mm. at the reduced air pressure.

(*b*) When the metal specimen is connected directly to the electroscope, as in the majority of the experiments, it is possible to test for a charge on the insulator specimen by discharging and isolating the metal specimen, and then putting it close to the insulator. If the insulator is charged, a charge is induced and the electroscope shows a proportional deflection. It is found that if a charge is generated in air, the specimens separated and the air evacuated, then the charge remains on the insulator.

This shows that the diminution in charge found earlier is not due to the removal from the surfaces of a layer of charged gas particles.

(*c*) Experiments with the metal specimen connected to the electroscope have shown that when the charged specimens are separated by even 1 mm., there is no decrease in the charge on reducing the air pressure. This is in contrast to (*a*) above, where the metal being isolated is at a high potential, and will discharge under these conditions.

(*d*) When the charged specimens are left in contact in air for periods up to 16 hours, there is no appreciable variation of charge.

Discussion of above Results.

The main points brought out in the above experiments are :—

(1) Only very small charges are obtained at low pressures of the surrounding gas.

(2) The large charges obtained in gases are found to disappear if the specimens are left *in contact* and the gas evacuated. If the specimens are separated by a small distance, this discharge does not occur.

(3) The metal specimen, when not connected directly to a large capacity (*e.g.*, the measuring system), is at a potential high enough to cause discharge to a surrounding metal cylinder when the air pressure is reduced.

(4) The charges obtained at pressures intermediate between atmospheric and the lowest pressure used are roughly proportional to the pressure.

From a consideration of these results, it appears that the charge obtained when two substances are rubbed together is limited in magnitude by the dielectric strength of the medium between.

In all cases of rubbing in air there is effectively a layer of air between the specimens, and it appears that this layer is necessary in order to keep the charges apart. The results of these experiments show that the maximum surface density of charge obtained with any specimens is approximately 20–21 E.S.U. per square centimetre. Other workers, notably Richards,* have also obtained maximum charges of 20–21 E.S.U. per square centimetre.

Now in the case of the present experiments, although the specimens have not been optically flat, it can be assumed without appreciable error that the average distance apart of the surfaces is very small. The question of the potential differences required to produce discharges over very short distances does not appear to be finalised. The results of Earhart, Shaw, Kinsley and Hobbs† show that discharges can be obtained with very small differences of potential, in some cases less than 1 volt. On the other hand, Almy‡ states that there is a minimum potential below which a discharge will not pass, no matter how short the distance. The results of the other workers, however, appear to be quite reliable and very similar to each other, and here it is assumed that their figures are correct. The surface densities of charge obtained in frictional electricity are sufficient to give the intensities found by these workers, and it is probable, therefore, that the intensity necessary to cause discharge is of paramount importance in determining the maximum frictional electric charge obtainable.

---

* 'Phys. Rev.,' vol. 22, p. 122 (1922).

† Earhart, 'Phil. Mag.,' vol. 1, p. 147 (1901); Shaw, 'Roy. Soc. Proc.,' A, vol. 73, p. 337 (1903); Kinsley, 'Phil. Mag.,' vol. 9, p. 692 (1905); Hobbs, 'Phil. Mag.,' vol. 10, p. 617 (1905).

‡ 'Phil. Mag.,' vol. 16, p. 456 (1908).

The conclusion arrived at from these results may then be stated :—

"The maximum charge obtained by friction between two bodies is limited in magnitude by the dielectric strength of the gas between the bodies. When the charges reach such a value that the resultant intensity is greater than the gas can withstand, then the charges come together and neutralise until they have fallen to a value lower than that necessary to give the breakdown intensity in the gas."

From this conclusion it follows that different charges will be obtained if the medium is changed or the dielectric strength of any one medium, say the air, is altered by ionisation, or some other means. The dielectric strengths of gases are all very similar, and hence we would not expect to obtain large variations with different gases. Liquids, however, have a different dielectric strength from air, and we would expect to obtain different charges if the specimens were surrounded by liquids instead of air.

Experiments have been made to test these effects.

The Effect of ionising the Interspace.

The γ-rays from radium were used to produce ionisation. These are sufficiently penetrating not to be appreciably reduced by the thickness of the specimens. The results of every experiment have shown that exposure to radium causes an immediate decrease in the charge, but the rate of decrease diminishes rapidly, and continued exposure produces a slow but steady decrease in the charge.

This slow decrease, due to an ionisation leak, is in sharp contrast with the rapid decrease which occurs when there is a discharge due to the breakdown of the dielectric strength of the air.

The specimens were rubbed in air to give the maximum charge, then put in contact, isolated, and a tube containing radium bromide placed about 2 cm. from the specimens (resting on the seat of the window in the cover, the plug being removed). After varied intervals, the radium was removed to a distance, and the charge on the specimens measured. The specimens were then put in contact again, and the radium replaced. It was found that if the radium was not removed while the specimens were separated to measure the charge, the sulphur specimen was almost completely discharged. This was due to the large ionisation produced in the air itself when the rays pass through. When the specimens are in contact, probably the majority of the ions are ejected from the specimens themselves. When the specimens were left in contact with the radium removed, no decrease of charge occurred.

The results of a typical experiment are shown in graphs 7 and 8. After rubbing to generate the charge, the specimens were left in contact for one hour. The radium was then brought up (Point A, graphs 7 and 8). The specimens were subjected to the radium for 80 minutes, the charges being measured at intervals (Point B). The radium was removed, and the specimens left in contact for 14 hours (Point C). Then the radium was applied again for 21 hours.

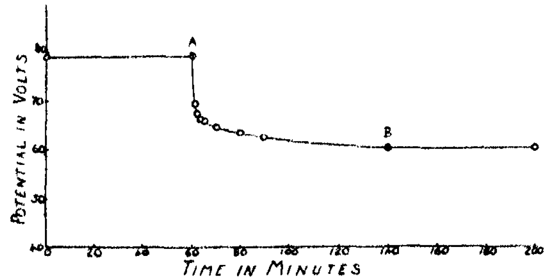


Fig. 7.—The Effect of Radium. Radium applied at A, removed at B.

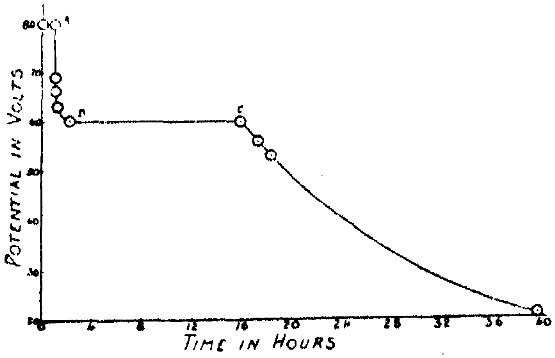


Fig. 8.—The Effect of Radium. Radium applied at A and C, removed at B.

Graph 7 shows the initial part of the experiment, and graph 8 all the values. The values are given as the potential of the measuring system.

The graphs show clearly the discharging action of the γ-rays.

This experiment is important as it shows that the charges are not maintained by an intrinsic difference of potential, but will re-combine when the air between is made conducting.

H. F. Richards (*loc. cit.*) obtained no effect on radiating his surfaces with X-rays, and hence concluded that the charges are maintained by a contact potential difference.

In the present investigations experiments were first conducted using X-rays to produce ionisation. Although a strong beam was used no effect was observed. The anticathode of the X-ray bulb was about 30 cm. from the specimens, and the height of the bulb would have to be adjusted with great accuracy in order to cause the X-rays to pass freely between the specimens when the latter are in contact. Attempts were made to adjust the bulb, but the probability is that, due to incorrect alignment, the X-rays were not passing into the interspace in sufficient amount to produce ionisation.

The Effect of a Liquid Medium.

The dielectric strength of oils is for ordinary distances many times greater than that of air, but Shaw and Earhart* have shown that for spark distances of less than $2 \cdot 5 \times 10^{-3}$ to 6×10^{-3} mm., varying with the liquid used, air has a dielectric strength greater than that of the liquids. They found air to be from 1·5 to 5 times as strong as the various liquids used.

It is to be expected, therefore, that when the specimens in frictional electricity are surrounded by an oil, a smaller charge will be obtained than in air.

Experiments have been made to test this, using petroleum jelly and paraffin oil. In each case a smaller charge has been obtained than in air.

I. *Petroleum Jelly.*—This was used in the purified form as supplied for the greasing of special apparatus. It was applied to the specimens in a thick coat by rubbing on with the finger.

A series of rubs were given in air, then the jelly smeared on and another series given.

The specimens were sulphur and aluminium.

Mean charge in air .. .. = 10·3 E.S.U./sq. cm.
Mean charge in jelly .. .. = 6·2 E.S.U./sq. cm.

The experiment was repeated on several occasions with very similar results.

These results show definitely that the presence of a layer of jelly largely reduces the charge obtained.

---

* Shaw, 'Phil. Mag.,' vol. 12, p. 317 (1906); Earhart, 'Phys. Rev.,' vol. 23, p. 358 (1906).

This is the effect observed by workers who have measured the charge produced when various substances are rubbed with textiles. After considerable rubbing the charge is often found to decrease suddenly and simultaneously a smear of grease from the textile appears on the surface of the specimen.

II. *Pure Paraffin Oil.*—Common paraffin oil was taken and re-distilled over sodium. A small cylindrical cup of sufficient size to allow the upper specimen to be rotated inside it was made of insulating material, and screwed on the support for the metal specimen. The specimen was then screwed inside the cup so that the oil could be poured in and would cover it. The method adopted was to give a series of rubs in air, then pour in the paraffin and give another series.

The results show a decided decrease in the charge obtained with the paraffin. The following result is typical:—

Specimens—aluminium and sulphur.

Average charge in air .. .. = 11·6 E.S.U./sq. cm.
Average charge in paraffin .. = 5·3 E.S.U./sq. cm.

From a graph (fig. 4, p. 324) in Shaw's paper we find that for a distance of 5×10^{-5} cm. the ratio of the potentials required to cause a discharge in air and in paraffin is almost 2·4 to 1, while the ratio of the charges obtained in the present experiments in air and paraffin is 2·2 to 1.

This, then, is strong evidence that the magnitudes of the different charges obtained are determined by the dielectric strength of the medium.

The Effect of Time of Contact after rubbing in Impure Paraffin.

In some experiments with impure paraffin oil there appeared to be a decrease in the charge if the specimens were left in contact after rubbing. An experiment was therefore made to test this.

A series of readings were taken separating the specimens while the top one was still rotating. This gave us the charges with zero time of contact. Other series were taken leaving the specimens in contact for different periods after the rub was completed before measuring the charge.

The values obtained are shown on the graph (fig. 9).

These results are instructive as they show that there is re-combination between the charges generated.

This result is similar to that obtained when the air is ionised by radium.

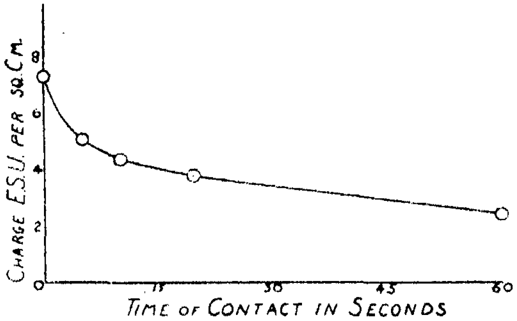


FIG. 9.—Decrease of Charge in Impure Paraffin with Time of Contact.

The Effect of substituting other Gases for Air.

In some experiments described earlier, different gases were used as the medium surrounding the specimens. These showed slight variations in the charge produced, but nothing beyond the usual differences obtained under apparently identical circumstances.

In order to find if there was any large difference caused by changing the gas, comparable with the effect of evacuating, the following procedure was adopted with specimens of aluminium and sulphur :—

(1) The specimens were given a series of 10 rubs in air and the charge obtained noted.

(2) The specimens were left in contact, isolated, and the gas passed in for 5–10 minutes through one inlet while the air was drawn out through the other opening by the pump.

(3) The specimens were separated and the charge measured. This would show any sudden change caused by introducing the gas.

(4) The specimens were put in contact, and 10 rubs given as in (1). The charges noted.

(5) Specimens left in contact, isolated, and air passed in for 5–10 minutes while the gas was drawn out by the pump.

(6) Specimens separated and charge measured.

(7) The whole procedure was then repeated two or three times.

All gases passed into the box were dried by passing through calcium chloride tubes and wash bottles of concentrated H_2SO_4.

The surfaces of the specimens were cleaned before each experiment.

I. *Sulphur Dioxide.*—The gas was obtained from a bottle of the liquid :—

Gas and time of passing.	Charge on testing.	Average charge on rubbing.	Specimens isolated at
Air	—	7·1	7·1
SO_2 (5 mins.)	7·1	7·2	7·2
Air (10 mins.)	7·2	6·9	7·1
SO_2 (5 mins.)	7·1	7·3	7·1
Air (5 mins.)	7·1	6·9	—

II. *Carbon Dioxide.*—The gas was obtained by the action of HCl on marble in a Kipps apparatus :—

Gas and time of passing.	Charge on testing.	Average charge on rubbing.	Specimens isolated at
Air	—	10·2	10·4
CO_2 (5 mins.)	9·6	10·3	9·5
Air (5 mins.)	9·5	10·4	10·4
CO_2 (45 mins.)	10·4	10·5	—

III. *Coal-Gas.*—This was obtained from the city supply :—

Gas and time of passing.	Charge on testing.	Average charge on rubbing.	Specimens isolated at
Air	—	9·5	9·5
Coal-gas (10 mins.)	9·5	9·5	9·5
,, (30 mins.)	9·5	9·5	9·5
,, (40 mins.)	8·5	9·5	9·5
Air	8·5	9·3	—

IV. *Hydrogen.*—The gas was prepared by the action of sulphuric acid on zinc in a Kipps apparatus. It was passed through solutions of $AgNO_3$ and $K_2Mn_2O_8$ in order to purify it :—

Gas and time of passing.	Charge on testing.	Average charge on rubbing.	Specimens isolated at
Air	—	8·5	8·5
H_2 (5 mins.)	8·5	6·9	6·4
Air (5 mins.)	6·4	6·7	6·9
H_2 (5 mins.)	6·9	6·4	6·9
Air (10 mins.)	6·9	6·8	—

V. *Oxygen.*—The oxygen was obtained from a cylinder of compressed gas, stated by the manufacturers to be 99·9 per cent. pure :—

Gas and time of passing.	Charge on testing.	Average charge on rubbing.	Specimens isolated at
Air	—	10·0	9·4
Oxygen (10 mins.)	9·2	9·5	9·4
Air (10 mins.)	9·4	10·2	10·5
Oxygen (10 mins.)	9·6	9·6	9·6
Air (5 mins.)	9·6	10·2	10·0
Oxygen (5 mins.)	9·6	10·1	10·0
,, (5 mins.)	9·6	9·4	9·1
Air (5 mins.)	8·7	9·6	9·1

The above results show that in the charges obtained with these gases there is no difference from air within the limits of accuracy of these experiments.

J. H. Jones* measured the charges produced on a metal specimen rubbed by a rotating wheel covered with silk. The metal was connected directly to an electroscope, and he took his readings with the specimen in contact with the wheel. After the charge reached a maximum, and while the wheel was running, he introduced other gases instead of the air, and found in all cases, except SO_2, that there was a decrease in the charge, sometimes small, more often fairly large.

With oxygen he found a decrease as large as 50 per cent. in his charge, and with nitrogen a decrease between 40 and 50 per cent. Carbon dioxide, the other main constituent of air, also gave decreases of 20–30 per cent. It is extremely unlikely that the constituents of a mixture such as air will show such markedly different properties from the mixture itself.

As is shown above, no large differences could be detected with various gases which had been carefully dried.

Jones makes no mention of drying his gases or of the sources from which he obtained them. If prepared by ordinary laboratory methods, they would be almost saturated with moisture, and the presence of this moisture would be sufficient to explain the decreases he obtained. Many workers have noticed that often no result is obtainable in damp air even when the insulation of the specimens is good, and the reason is that the damp air allows the charges to re-combine more readily than dry air.

This may be the reason for Jones' results with gases.

---

* 'Phil. Mag.,' vol. 50, p. 1160 (1925).

The Effect of the Capacity of the Specimens.

(*a*) In all experiments carried out with the metal specimen connected directly to the electroscope, we found that there was no deflection of the electroscope while the specimens remained in contact after rubbing. Immediately on separating them, however, the electroscope gave a sudden deflection, which then increased slowly to a maximum when the specimens were separated by a distance of 2–3 cm.

If, now, the specimens were put together again without discharging, the electroscope deflection decreased as they approached one another, and dropped to zero as soon as they were in contact again.

This showed that during the rubbing all the charge produced remained on the interface, none spread to the rest of the system, and hence the charge produced was independent of the capacity of the holder of the metal specimen.

This was shown clearly by the fact that the same charge was obtained using method 1 when the metal specimen was isolated and had a capacity of 3–4 cm., and in method 2 where the specimen was connected to the measuring system and thus had a capacity, when isolated, of almost 80 cm.

It was also shown by the fact that having the metal specimen connected to earth—*i.e.*, of infinite capacity during the rubbing—makes no difference to the charge produced. The connection to earth has, of course, to be broken before the specimens are separated.

(*b*) In a certain experiment with steel against glass, using method 3, it was found that the charge obtained on the steel apparently kept on increasing until the limit of the high-tension battery used for comparison was reached. This was found to occur for several days even when the specimens were cleaned and washed.

The explanation was found when it was noticed that even when the specimens were in contact the electroscope showed a deflection. This deflection was found to increase between each rub by the same amount as the total charge increased.

This meant that each time the specimens were separated, part of the charge on the insulator leaked away, and when the specimens were placed in contact again, all the charge on the metal specimen was not drawn to the interface. Part remained on the rest of the system, and the electroscope therefore showed a small "residual" deflection. The next rub would bring the charge on the interface up to the previous value, and then, on separating, the charge on the metal spread to the rest of the system, giving a reading which was larger than the previous reading by the amount of the residual charge. On putting the specimens in contact again, if the same amount had leaked from the insulator, the residual charge on the electroscope was greater than before, and thus the process continued, the residual charge increasing step by step with the main charge. The type of the effect is shown by the following figures for an experiment with steel and glass. The values in the top line show the potentials obtained after each rub, using method 3 ; those in the lower line the corresponding "residual" electroscope potential, *i.e.*, the potential shown before separating the specimens. Thus, the first rub gave a potential of 182 on the measuring system, there being no deflection till the specimens were separated. Then, on putting in contact again, the electrometer still showed a charge of 60 volts. Another rub was given and then the charge obtained was 245 on the measuring system, which is almost 182 + 60. Similarly through the series.

Voltage of measuring system	182	245	414	442	552	646	704	712	734
Residual voltage	—	60	220	260	380	455	515	520	560

The values are only approximate, particularly those for the residual voltage, which were not read very accurately. The electroscope was far too sensitive to read these high charges, and an opposing charge of 400 volts had to be put on the plate.

The fact that the increase between the different steps varies is due to the specimens being separated for different periods each time, thus allowing different amounts to leak from the insulator. For instance, there was a longer interval between the second and third readings, as the potential on the divider had to be increased to measure the higher potentials.

Similar results were obtained on another occasion with aluminium and sulphur, when we obtained with method 3 :—

First rub	94 volts.
After 25 rubs	158 volts.
Residual electroscope potential	61 volts.

The leakage from the upper specimen was found to be due in each case to a smear of grease accidentally put on the insulating support while passing the upper holder through the upper taper fitting. After cleaning the insulation, there was no trace of this effect.

Thus it is seen that when the charge is able to leak away from the dielectric specimen, the total charge on the metal specimen and its connected system keeps on increasing, and the maximum charge obtainable will depend upon the capacity of the metal specimen and its support. These conditions prevail when a charge is generated on a metal by rubbing with a semi-conducting wheel.

The Maximum Charge obtainable under Various Conditions.

Assuming that the magnitude of the charge is limited by the dielectric strength of the medium, we can discuss the charges obtainable under various experimental conditions:—

(*a*) With specimens rubbed as described in the present paper. In this case all the charges remain on the surfaces, and will increase until the intensity between the opposite charges breaks down the dielectric between. In this case we are dealing with the simplest conditions.

(*b*) The other method of rubbing the specimens is that employed by Owen, Morris Jones, French, Cardani and various others.* In each case these workers had a small specimen, 2–3 cm. diameter, held against the rim of a wheel, in some cases 3 feet in diameter, the wheel being in some cases conducting, in others insulating.

(1) Consider first the case where the specimen is conducting and connected to a capacity, and the wheel is a perfect insulator and any charge generated on it remains in its initial position. During the first revolution of the wheel fresh portions of the rubber without any charge are brought into contact with the charged specimen. The result will be similar to that described above, when charge could leak from the insulator specimen—*i.e.*, in this case part of the charge on the surface of the specimen will spread to the rest of the specimen system.

During the second revolution each portion of the wheel will already have a charge upon it due to the first revolution, and if it is assumed that the generation of charge is not affected by the presence of earlier charges, at the end of the second revolution each portion of the wheel will have twice the charge it had at the beginning of the revolution. A charge equal to this will be on the surface of the specimen, and the surplus charge generated on the specimen during the rub can spread to the rest of its system.

Thus the process goes on until the charges on the wheel and the surface of the specimen are sufficiently large to cause discharge.

(2) When the wheel is insulating and the specimen has little capacity other than its surface—*i.e.*, in the extreme case the specimen is an

---

* *Loc. cit.*

insulator—then there is an accumulation of charge on the specimen surface, and discharge will take place after a smaller number of rubs than in case (1) above.

The total charge generated will therefore be less—that is, the total charge obtainable when the wheel is an insulator depends upon the capacity of the specimen system.

(3) A third case is when the specimen is a conductor and the wheel a partial conductor and connected to earth. As soon as each portion of the wheel is removed from contact with the specimen, part of the charge on this area will leak away. Thus, at the end of each revolution there will be less of the charge held on the surface of the specimen by the charge on the wheel opposite—*i.e.*, a larger percentage is free to spread to the rest of the specimen system. Thus, if, as in the first case, the capacity of the specimen system is large, then the charge actually on the surface of the specimen will be less at any instant than at the same instant if the wheel is an insulator—*i.e.*, discharge will not take place as soon as when the wheel is an insulator, and therefore a larger total charge will result. Any re-combination of the charges due to actual conduction is equivalent to decreasing the charge generated by each revolution, and will merely increase the rubbing necessary to give the maximum charge, which is unaltered.

In the extreme case where both specimen and wheel are perfect conductors, no charge can be obtained.

(4) When the wheel is a conductor and the specimen an insulator, all the charge generated on the specimen must remain on its surface, and on the conducting wheel an equal and opposite charge is induced. Thus discharge will occur sooner than when the capacity of the specimen is greater—*i.e.*, a smaller charge will be obtained when the specimen is an insulator.

We have, then, that when the specimen is a conductor, the maximum charge obtained is

(1) Proportional to the capacity of the specimen system.
(2) Greater the greater the conductivity of the wheel.

These are the experimental results found with this type of apparatus by Morris Jones, J. H. Jones and the other workers. By making several assumptions, it is possible to calculate mathematically the maximum charges obtainable under the various conditions. The results agree entirely with the discussion above, but are not given here, since the only main principles regulating the maximum charge are under consideration.

From the above results it follows that to obtain the largest charge in a frictional electric machine such as was used by early workers, the best arrangement is with a specimen conducting and connected to a large capacity, while the rubber is completely discharged between the successive rubs.

The Relation between Contact Electricity and Frictional Electricity.

By contact electricity is meant the charge which is obtained when two specimens are placed in contact without any motion in the plane of contact. Frictional electricity is distinguished as the charge obtained when, after being placed in contact, one specimen is rotated or otherwise moved in the plane of contact.

Helmholtz's theory states that frictional electricity is merely contact electricity, the rubbing being necessary to bring every portion of the surface of each specimen in contact with the surface of the opposite specimen. It follows from this, as pointed out by Owen, Morris Jones, Richards and others, that with two specimens absolutely flat, so that on placing them together they are in contact all over their surfaces, then as large a charge should be obtained with mere contact as by rubbing the two together.

Richards found with his optically flat surfaces that contact, even under a pressure of 75 kgm. per square centimetre, did not give him more than 15 per cent. of his maximum charge, which was obtained with a rotation of ¼ revolution. He concludes, however, that Helmholtz's theory is correct, since ¼ revolution gave his maximum charge.

If, however, the frictional charge is due to, say, mechanical disruption of the molecules, then, as the relative movement of the specimens in a ¼ revolution is very large with reference to the diameter of a molecule, it may well be that this movement is sufficient to give the maximum charge that the air or other medium can withstand.

Richards does not state what precautions, if any, he took to see that in placing his specimens together he did not have a slight rotation. It will be seen that when ¼ revolution gives his maximum charge, a very slight movement is sufficient to give 15 per cent. of the maximum. This is shown by the shape of the curves, as in graph I. All other workers have found that the charges on mere contact are extremely small.

J. H. Jones squeezed platinum and silk together in a vice and obtained only one-twentieth of the charge given by a very small amount of rubbing. In the present research it was found that pure contact gave only from 0·3 to 2·5 per cent. of the charge obtained by even ¼ revolution.

It has been shown in the experiments above that an insulating medium is necessary between the specimens in order to obtain a measurable charge. This means that the charging produces an unstable condition, and is therefore probably due to some disruption of the molecules of one or both of the specimens. Possibly electrons are freed from their atoms during the rubbing and accumulate on one specimen. When the rubbing ceases we have two opposite charges of large surface density separated by a thin layer of dielectric. If the dielectric is weak enough or the charges strong enough, they will neutralise each other, and there is no resultant charge on either specimen.

If, now, we have perfectly flat specimens in contact all over their surfaces, there cannot be any insulating medium between. On rubbing the specimens together, therefore, we may have the separation of charges as before, but when rubbing ceases there is nothing to keep the charges apart, and hence they re-combine—*i.e.*, with perfectly flat surfaces it would be impossible to obtain frictional electric charges.

It appears from this that the separation of frictional electric charges is not due entirely to pure contact between the surfaces, or, in other words, that Helmholtz's theory is not correct.

Discussion of Coehn's Law.

The law enunciated by Coehn may be expressed by $Q = C(K_1 - K_2)$, where Q is the charge produced per unit area on a substance of dielectric constant K_1 when rubbed by another with dielectric constant K_2; C is a constant. It is proposed to discuss several results of other workers which bear on this law.

Morris Jones found that in every case he obtained a larger charge when he rubbed his specimens with silk than with chamois leather, although some received a negative and some a positive charge. The sign of charge on any particular specimen was in every case the same with both materials. It is obvious that the magnitude of these charges does not depend upon the relative positions of the substances in the frictional electric series, since if the substances which receive a positive charge have their larger charge with silk, then those which receive a negative charge should receive their larger charge with chamois. In these results, therefore, Coehn's law does not apply. From the results of the present paper it appears probable that the wax, which is present in the chamois much more than in silk, spreads over the surfaces and allows discharge at a lower potential than can occur with the silk.

Richards assumed the above equation to be correct, and from four experiments with dielectrics worked out the constant, which he found to vary by 30 per cent., namely, from 3·90 to 5·06. Other experiments with steel and dielectrics showed that if he gave the value 3·1 to the dielectric constant of steel, the constant C came between these limits. He therefore concludes that Coehn's law is true, and that the dielectric constant of steel is 3·1.

This value for steel, however, requires a constant of 62 for steel rubbed on ebonite, and Richards found also that he obtained charges with specimens of the same material. This last result has been shown by many other workers, and also in papers by Shaw and Jex* it has been shown that the sign on a particular substance can be varied at will by physical processes which will not affect the dielectric constant. Considering these facts, it is apparent that Coehn's law is not yet proved for solid specimens.

It is highly probable that in most cases the maximum charge obtained is determined by the dielectric strength of the medium (the air in most cases), and the various substances influence the maximum charge only in so far as their surface form and condition alters the amount of the medium between.

Results obtained by Richards and by J. H. Jones.

Richards measured the charges produced on rubber when compressed against various substances. He found, using two forms of rubber of different dielectric constants against various insulators of widely differing dielectric constants, and also against lead and steel, that in every case he obtained maximum charges of the same sign, and although the magnitude varied slightly, in every case it approximated to the maximum values obtained in frictional electricity, *i.e.*, from 15–21 E.S.U. per square centimetre. He found the charges increased with pressure up to 2 kg. per cm.^{2}, and then remained constant with pressures up to 5 kg. per cm.^{2}. It is probable that in each case he was obtaining the maximum charge the surrounding medium (air) would withstand.

J. H. Jones used wheels of various materials as base for his silk rubber, and took all his readings with the metal specimen in contact with the wheel. Thus, as was shown earlier, his electrometer deflection will be proportional not to the charge on the specimen and its connected system, but to the difference between the total charge generated on the specimen and the charge on the area of the wheel immediately opposite the specimen. The charge on the wheel will depend upon its conductivity, being smaller the higher its conductivity.

---

* Shaw and Jex, 'Roy. Soc. Proc.,' A, vol. 111, p. 339 (1926); Shaw, 'Proc. Phys. Soc.,' vol. 39, p. 449 (1927).

Thus, with a good conductor, larger electrometer deflections should be obtained. This he found to be the case. In his experiments he found that after some rubbing his deflection reached a maximum. After this stage there were sudden drops to smaller deflections, followed by a more gradual rise to the previous maximum. From the results discussed earlier in the present paper it is suggested that these decreases are caused by discharge from specimen to wheel when the charges are sufficient to break down the air between. Part of the charge disappears, and then the continued rubbing brings it up to its previous value, and the discharge occurs again.

Summary of Results and Conclusions.

1. The maximum charge obtainable by friction between any two substances is independent of

(*a*) The normal pressure between the specimens.

(*b*) The velocity of rubbing.

(*c*) The amount of rubbing, provided this is greater than a certain minimum necessary to give the maximum charge.

2. In no case, with even the flattest specimens, has a charge been obtained on mere contact at all comparable in magnitude with that given by even ¼ revolution.

3. The magnitude of the maximum charge obtained between any two specimens depends largely on the dielectric strength of the medium between the specimens. When the dielectric strength is high, larger charges are obtained than when it is low.

The maximum charges between any two substances in air are not, then, necessarily characteristic of the two substances, but may be merely an indication of the amount of air held between the surfaces, due to adsorption or other surface effects.

4. The large variations found by many experimenters in the maximum charges between the same substances are readily explained on the basis of slightly different surface conditions, giving discharge with different surface densities of charge.

This also explains the irregular variations found in individual experiments.

5. The charges are not maintained by an intrinsic difference of potential between the substances, as was stated by Richards.

6. The charges are not due entirely to contact between the specimens, but rotation is necessary to obtain the maximum charge—*i.e.*, Helmholtz's theory is not correct.

7. The various common gases are found to have the same effect as air, and the same charges are obtained in each one.

8. Experiments on the generation of charges with a rotating wheel are very indefinite, and many more factors are involved than with the apparatus used in the present investigation.

In conclusion, the writer desires to thank Prof. Burbidge for suggesting the subject of frictional electricity, and for his interest and advice throughout the course of the work. He is also indebted to Mr. G. McB. Salt for taking the various photographs.

The Structure of Cyanite, Al_2SiO_5.

By W. H. TAYLOR and W. W. JACKSON, Manchester University.

(Communicated by W. L. Bragg, F.R.S.—Received March 7, 1928.)

Introduction.

The three naturally occurring forms of Al_2SiO_5 have been examined by Mark and Rosbaud,[*] who have determined the space-groups of the orthorhombic forms (andalusite and sillimanite) and the unit cell of cyanite, which belongs to the pinacoidal class of the triclinic system. The present paper contains an account of the determination of the arrangement of the atoms in cyanite.

The analysis of this structure is interesting from two points of view. In the first place, it is of special interest to the mineralogist because it is probably the most striking example of a crystal which exhibits different degrees of hardness in different directions on the same face. The measured values of the hardness on the (100) face of cyanite are 4–5 in a direction parallel to the *c* axis and 6–7 in a direction parallel to the edge common to the (100) and (001) faces.[†] Any suggested structure must, of course, explain this unusually large variation.

Secondly, the determination of this structure serves to show how certain general guiding principles may be employed in the analysis of crystal structures by X-ray methods in order to simplify the solution of a problem so complex that a direct attack would be impracticable. The possibility of analysing this crystal was suggested when it was recognised that the structure might be based

[*] 'N. Jahrbuch f. Min.,' Beilageband 54, A, p. 127 (1926).

[†] Dana, 'Mineralogy,' p. 500.

on an assemblage of oxygen atoms arranged in one of the forms of closest packing; the only other assumption which we have found it necessary to make is that the silicon and aluminium atoms fit in the interstices between the oxygen atoms in the same way in this crystal as in all other silicates which have been analysed. The use of such guiding principles in the analysis of a number of complex silicates has already been discussed very fully*; their application to cyanite will be described in detail in a later section.

Description of Method.

The determination of the structure is qualitative in nature, in the sense that no absolute measurements of intensity of reflection have been made; we have estimated the relative intensities of a large number of reflections obtained in a number of single-crystal rotation photographs. Special care has been taken to make the estimated intensities from different photographs comparable, to choose suitable crystal specimens, and to make due allowance for the various factors which influence the density of the photographic image produced by a reflection of given intensity when the rotating-crystal method is employed.

We have assumed the scattering power (F-curves) of the silicon atom and the aluminium atom to be identical for all glancing angles; this is true within the accuracy of our experimental observations. Absolute F values have not been used, but a rapid decrease in scattering power, with increasing glancing angle, has been assumed; the decrease has been assumed to be more rapid for the oxygen atom than for the silicon and aluminium atoms, in accordance with the experimental results of Claassen.†

In an accurate quantitative investigation of the structure, using absolute measurements of intensity of reflection from a relatively small number of planes, it would be necessary to know the degree of perfection of the crystal, in order to apply corrections for primary and secondary extinction, and to know how to interpret the observed intensities in terms of the F values of the atoms concerned.

In our qualitative method we simply seek a general agreement between the way in which the observed intensity varies for different reflections and the variation of the atomic contributions to the calculated structure-factor; it is therefore not necessary to correct for extinction, nor to know the degree of perfection of the crystal. The amount of data available is large by comparison with that usually obtained in quantitative work, and we consider that this amply compensates the inaccuracies inherent in a qualitative investigation such as the present one.

---

* W. L. Bragg and J. West, 'Roy. Soc. Proc.,' A, vol. 114, p. 450 (1927).

† 'Proc. Phys. Soc.,' vol. 38, p. 482 (1926).

Proceeding in this way, a structure has been found which gives a general agreement between the calculated structure-factors and the observed intensities of reflection for a large number of planes.

*The Arrangement of the Oxygen Atoms.**

Crystallographic data show that cyanite belongs to the pinacoidal class of the triclinic system, and has the following axial ratios and angles :—

$$a : b : c = 0{\cdot}8894 : 1 : 0{\cdot}7090 \quad \alpha = 90^\circ\ 5\tfrac{1}{2}' \quad \beta = 101^\circ\ 2' \quad \gamma = 105^\circ\ 44\tfrac{1}{2}'.$$

Careful measurements of the axial lengths, obtained from rotation photographs, and checked by spacing measurements on the ionisation spectrometer† lead to the following values‡ :—

$$a = 7{\cdot}09 \pm 0{\cdot}02\ \text{Å}. \qquad b = 7{\cdot}72 \pm 0{\cdot}04\ \text{Å}. \qquad c = 5{\cdot}56 \pm 0{\cdot}02\ \text{Å}.$$

The unit cell formed by this set of axes contains four molecules of Al_2SiO_5.

The refractive index of cyanite (1·72), and the volume occupied per oxygen atom in the crystal (15·05 Å.^{3}), indicate that its structure may be based on one of the close-packed arrangements of oxygen atoms, which would have refractive

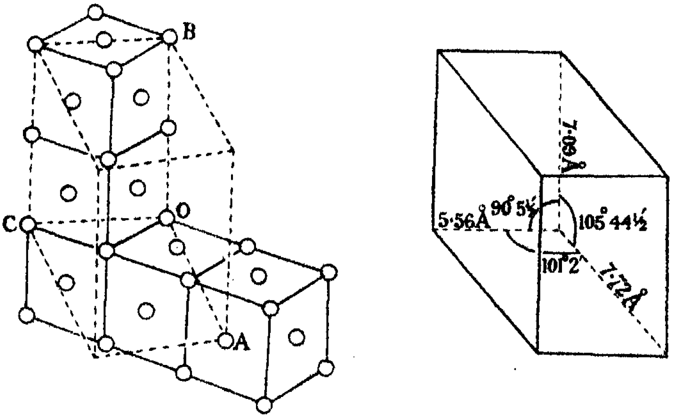


Fig. 1.—Relationship between the Triclinic Cell derived from the Cubic Lattice (left) and the Unit Cell of Cyanite (right).

---

* For detailed discussion of contents of this section, *vide* Bragg and West, *loc. cit.*, pp. 457–462.

† For the spacing measurements we are indebted to Mr. J. West.

‡ Mark and Rosbaud give the values $a = 7{\cdot}18$ Å., $b = 8{\cdot}00$ Å., $c = 5{\cdot}55$ Å. The value of c is easily obtained with considerable accuracy, and their value agrees with ours. The values we obtain for a and b are the result of several independent and highly consistent measurements.

index 1·71 and a volume per oxygen atom of 13·94 Å.^{3}. Trial shows that in a cubic close-packed arrangement a set of axes may be chosen, which show axial ratios and angles corresponding to those of cyanite; for in the diagram (fig. 1)

$$OA : OB : OC = 0\cdot935 : 1 : 0\cdot707 \quad \alpha = 90^\circ \quad \beta = 100^\circ\, 53' \quad \gamma = 105^\circ\, 38'$$

values nearly identical with those quoted above for cyanite. If we further assume a diameter of 2·70 Å. for the oxygen ion, the lengths of the axes in the close-packed array are:—

$$OA = 7\cdot14 \text{ Å.} \qquad OB = 7\cdot64 \text{ Å.} \qquad OC = 5\cdot40 \text{ Å.}$$

and the unit cell outlined by these axes contains 20 points of the lattice, corresponding to 20 oxygen atoms in the unit cell.

A simple experimental test of this suggested arrangement of oyxgens is possible if we assume that the silicon and aluminium atoms lie at the centre of groups of four and six oxygen atoms respectively; these groupings of oxygens round silicon and aluminium atoms have been found in every silicate which has previously been analysed, so that it is very probable that they will occur in this crystal also. This assumption is sufficient to show that planes of the types (200), (220), (222), (400) (referred to the cubic lattice) should give powerful reflections, although distortion of the structure will modify slightly the atomic contributions to the reflections from different planes. Rotation photographs were taken with the *b* axis and *c* axis of cyanite as axes of rotation, which corresponds to rotating the cubic lattice round its edge, and round its face-diagonal, respectively; the diagram (fig. 2) shows that the most intense spots correspond to the cubic planes specified above.

This evidence seems to show definitely that the triclinic crystal is based on a simple cubic arrangement of close-packed oxygen atoms, and supports the assumption that the silicon and aluminium atoms lie at the centre of groups of four and six oxygen atoms respectively. The existence of distortion from this ideal arrangement, and its amount, are discussed later.

The Method of determining the Structure.

Since cyanite belongs to the holohedral or pinacoidal class of the triclinic system, the space-group is C^1_i, and taking the origin of co-ordinates at the centre of the triclinic unit cell having the dimensions given in the previous section, the only independent symmetry elements are eight symmetry centres at the positions (000), ($\frac{1}{2}00$), ($\frac{1}{2}0\frac{1}{2}$), ($00\frac{1}{2}$), ($0\frac{1}{2}0$), ($\frac{1}{2}\frac{1}{2}0$), ($\frac{1}{2}\frac{1}{2}\frac{1}{2}$), ($0\frac{1}{2}\frac{1}{2}$). This gives

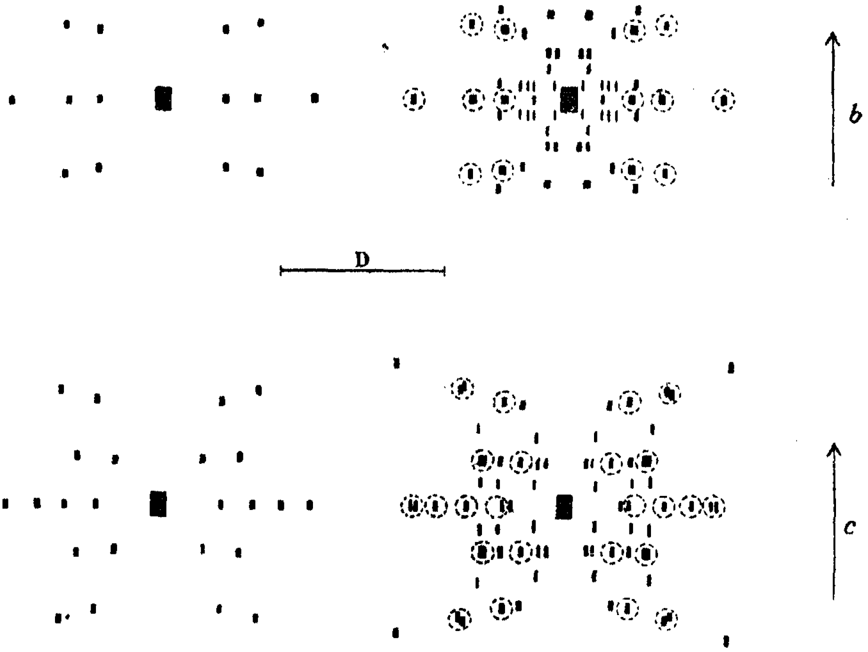


Fig. 2.—Comparison of Spots from Rotation Photographs of Cyanite (right) with those due to Principal Planes of the Cubic Lattice (left). Spots in the right-hand diagram corresponding to spots in the left-hand diagram are contained in circles. The length D represents the distance of the axis of rotation from the photographic plate.

a unit cell with symmetry centres at every corner, at the centre of every face, at the mid-point of every edge, and at the centre of the cell. An atom placed in any position, unless it lies at a symmetry centre, is converted into two atoms in the unit cell, by the operation of the symmetry elements.

We have to find the positions of eight aluminium atoms, four silicon atoms, and 20 oxygen atoms, since the unit cell contains four molecules of Al_2SiO_5. In assigning positions to these atoms, we make the following assumptions:—

(i) The oxygen atoms are arranged in cubic close packing.

(ii) Each silicon atom lies at the centre of a group of four oxygen atoms whose centres lie at the corners of a regular tetrahedron.

(iii) Each aluminium atom lies at the centre of an octahedron having an oxygen atom at each of the six corners.

We have assumed that the oxygen atoms are arranged accurately in cubic close packing, and in this case, using the customary values for the diameter of the silicon and aluminium ions, it is easy to show that these ions will just fit in the interstices of the oxygen assemblage, and so will lie at the centres of the holes occupied.

Trial shows that there are only two different ways of fitting the close-packed arrangement of oxygen atoms in the triclinic unit cell so as to satisfy the symmetry requirements of the space-group.

(*a*) In the first, oxygen atoms lie at four of the eight symmetry centres, viz., $(\frac{1}{2}00)$, $(\frac{1}{2}0\frac{1}{2})$, $(\frac{1}{2}\frac{1}{2}0)$, $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, and two oxygen atoms are in contact at each of the other four symmetry centres. Thus neither silicon nor aluminium atoms can lie at symmetry centres, and we have, in the unit cell, four oxygen atoms at symmetry centres, and eight pairs of oxygen atoms, two pairs of silicon atoms, and four pairs of aluminium atoms, in general positions.

This arrangement of oxygen atoms is illustrated in fig. 1, in which oxygen atoms lie at all the corners of the cell and also at the mid-points of the edges parallel to the *c* axis and the *b* axis, and the centres of the *a* faces of the cell.

(*b*) In the second, no symmetry centres are occupied by oxygen atoms, but two oxygen atoms are in contact at each of four of the symmetry centres. The four unoccupied symmetry centres are found to be at the centres of octahedral groups of six oxygens, and so may be occupied by aluminium atoms. In this arrangement, therefore, we have ten pairs of oxygen atoms and two pairs of silicon atoms in general positions and *either* four pairs of aluminium atoms in the general position, *or* three pairs in the general position with two at symmetry centres, *or* two pairs in the general position with four at symmetry centres.

This arrangement of oxygen atoms is obtained from the first, by giving the triclinic unit cell uniform translations of amounts $+ a/10$, $+ 3b/20$, $- c/10$ in directions parallel to the triclinic axes *a*, *b*, *c* respectively.

The number of ways of placing the silicon and aluminium atoms in the interstices of the oxygen arrangement in accordance with conditions (ii) and (iii) above is very large; for example, in either oxygen arrangement there are 24 "holes," any two of which may be occupied by the two independent silicon atoms. There are a correspondingly large number of ways of fitting in the aluminium atoms, but we have found that, of all possible ways of filling the holes so that the unit cell contains four silicon atoms and eight aluminium atoms (considering both oxygen arrangements), only one will account for the relative strengths of the reflections, from planes of low indices, observed on the rotation photographs.

The structure determined in this way is only approximate : it is certain that some distortion will occur, resulting in a movement of the silicon and aluminium atoms from the "ideal" positions which they have been supposed to occupy up to the present. If by the introduction of a set of small displacements from the "ideal" positions chosen for these atoms a much better agreement is obtained between observed and calculated intensities, this will serve as a strong confirmation of the essential correctness of the structure chosen. Such a set of small displacements is described in the next section. It is important to remember, however, that movements of the silicon and aluminium atoms will undoubtedly involve a corresponding distortion of the oxygen atoms : we have neglected this in our calculations, assuming that only those planes of the triclinic lattice which correspond to planes of simple indices in the cubic lattice will receive any contribution due to oxygen atoms. This is justified because planes of low indices are not seriously affected by slight distortion of the structure, while planes of high indices receive a relatively small contribution from the oxygen atoms owing to the more rapid decrease of scattering power with increasing glancing angle, in the case of oxygen as compared with silicon and aluminium.

Details of the Determination of the Structure.

(i) *Ideal Structure.*—The structure has been found by considering all possible arrangements of atoms which satisfy the conditions laid down in the previous section, and eliminating incorrect structures by a consideration of reflections from planes having simple indices ; a slight distortion from any ideal arrangement will have little effect on such reflections. To simplify this process of elimination in the initial stages, we have classified the possible positions for silicon and aluminium atoms in groups of four and two respectively, the atoms comprising each group being superposed when the structure is viewed in a direction parallel to the *b* axis (*cf.* fig. 5), so that they have the same co-ordinates parallel to the *a* and *c* axes, but different *b* co-ordinates. Consideration of simple reflections with indices $(h\,0\,0)$, $(0\,0\,l)$, $(h\,0\,l)$, which do not involve the *b* co-ordinate (*e.g.*, 100, 200, 001, 002, $20\bar{1}$, $20\bar{2}$), shows that there are only two possible ways of choosing silicon and aluminium groups for each arrangement of oxygen atoms ((*a*) and (*b*) in previous section)—*i.e.*, in all four ways of choosing silicon and aluminium groups.

Each such arrangement of groups, however, includes a large number of possible structures, since each silicon group represents four possible values for the *b* co-ordinate of a silicon atom, and similarly in the case of the aluminium atoms. Comparison of the observed intensities with the calculated structure-factor for the planes 010, 020, 030, reduces the large number of possible structures to seven, and by considering simple reflections such as 111, 110, $11\bar{1}$, 011, 012, 021, 211, $03\bar{1}$, these are further reduced to two, which we may call A_1 (based on the first oxygen arrangement (*a*)) and A_2 (based on the second oxygen arrangement (*b*)).

It is more difficult to decide which of these two remaining structures is correct, because both give structure factors in satisfactory agreement with the observed relative intensities of a large number of reflections from planes with simple indices. There are, however, a number of planes for which the observed intensity is very small and the structure-factor for A_1 very large (or vice versa), while no such disagreement occurs for the structure A_2. We are therefore justified in choosing A_2 as the correct " ideal " structure ; the improved agreement between calculated and observed intensities resulting from slight displacements from the atomic positions in this ideal structure will be described in the following paragraph ; it serves to confirm the structure found. The co-ordinates of the atoms in the ideal structure are given in Table I.

(ii) *Distortion from Ideal Structure.*—By choosing planes of suitable indices it is possible to determine the displacements of the atoms from the ideal positions, in directions parallel to *a*, *b*, *c* independently. (The angular co-ordinates $\theta_1 = 2\pi x/a$, $\theta_2 = 2\pi y/b$, $\theta_3 = 2\pi z/c$ were, of course, used in all structure-factor calculations.)

For the θ_1 co-ordinate (displacement parallel to *a*), the following series of planes were considered : 100 to 700, 102 to 602, $\bar{2}02$ to $\bar{7}02$, 111 to 611, $21\bar{1}$ to $71\bar{1}$, $\bar{4}11$ to $\bar{7}11$. It is permissible to use the ideal values of θ_2 and θ_3, while finding the slight variation from the ideal θ_1, because small variations of θ_2 and θ_3 will have little effect on the structure factor $\cos(h\theta_1 + k\theta_2 + l\theta_3)$ for the planes given above, while a small variation in θ_1 may have a big effect since *h* may be fairly large.

Similarly modified values of θ_2 and θ_3 have been obtained. The final co-ordinates of the atoms are given in Table I.

Table I.—Co-ordinates of Atoms in Cyanite.

Atom.	Number.	Ideal values.			Final values.		
		x/a.	y/b.	z/c.	x/a.	y/b.	z/c.
O_A	2	0·40	0·35	−0·40	Identical with ideal values.		
O_B	2	0·40	−0·15	−0·40			
O_C	2	0·20	0·05	−0·20			
O_D	2	0·20	−0·45	−0·20			
O_E	2	0·40	0·35	0·10			
O_F	2	0·40	−0·15	0·10			
O_G	2	0·20	0·05	0·30			
O_H	2	0·20	−0·45	0·30			
O_K	2	0	0·25	0·50			
O_M	2	0	0·25	0			
Si_I	2	0·20	0·42	−0·45	0·20	0·42	−0·48
Si_{II}	2	0·20	0·18	0·05	0·20	0·18	0·02
Al_I	2	0·20	−0·20	0·30	0·23	−0·21	0·30
Al_{II}	2	0·20	−0·20	−0·20	0·17	−0·21	−0·20
Al_{III}	2	0·40	0·10	0·10	0·38	0·11	0·07
Al_{IV}	2	0·40	−0·40	−0·40	0·38	−0·42	−0·43

The angular co-ordinates used in calculating structure-factors are given by the expressions $\theta_1 = 2\pi x/a$, $\theta_2 = 2\pi y/b$, $\theta_3 = 2\pi z/c$.

Using these co-ordinates the structure-factors for about 100 planes have been calculated, and the agreement between observed and calculated intensities is found to be much better, on the whole, than if the "ideal" co-ordinates are used. The results for about 40 planes are collected in Table II. The intensities quoted are estimated on an arbitrary scale having a maximum 10. The values of the glancing angle correspond to the K_α radiation from a molybdenum target ($\lambda = 0{\cdot}710$ Å.) and are only approximate; they are determined directly from the photographs by a simple graphical method, and are sufficiently accurate to enable an allowance to be made for the falling-off of the scattering powers of the various atoms with increasing glancing angle. It is, of course, essential to take this factor into account in comparing the observed relative intensities with the calculated structure-factors. In examining the results quoted in this table, it is also important to bear in mind that numerical agreement for individual reflections should not be sought; with the data at our disposal only a general correspondence in the variation of observed intensity and calculated structure-factor, with varying glancing angle, is to be expected. On the whole the agreement is good.

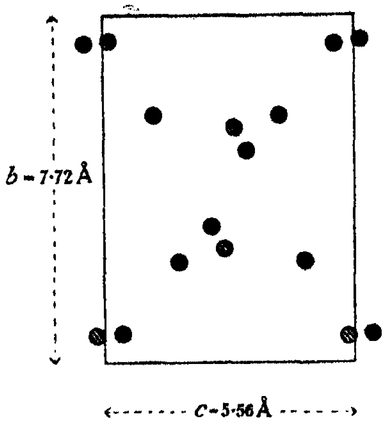
Table II.—Comparison of Structure-Factors and Observed Intensities.

Indices.	Glancing angle.	Observed intensity.	Structure-factor.
	°		
010	3	nil	0·1
100	3	nil	0·5
011	5	3	4·5
020	6	2	2·2
$\bar{1}20$	6	2	1·4
200	6	4	5·8
021	7	4	3·7
$\bar{1}2\bar{1}$	7	4	3·8
030	8	5	6·0
012	8	4	4·8
102	9	3	2·6
300	9	2	3·6
031	9	3	3·8
130	10	5	5·0
$1\bar{4}0$	10	7	23·6*
202	11	8	23·4*
040	11	2	1·7
013	12	1	0·5
140	12	2	1·3
400	12	nil	1·8
041	12	nil	0·3
302	13	nil	2·8
$1\bar{5}0$	13	nil	0·1
050	14	4	6·4
051	15	nil	3·0
150	15	3	4·2
$1\bar{6}0$	16	1	2·8
402	16	nil	2·0
500	16	7	30·0*
060	17	? trace	3·4
061	17	2	5·3
$1\bar{7}0$	19	2	1·6
502	19	2	5·4
600	19	nil	0·6
070	21	? trace	0·4
602	22	2	3·0
700	22	3	5·9
080	25	3	4·1

* The oxygen atoms make a large contribution to the structure-factors for these planes, which have simple indices when referred to the underlying cubic lattice of oxygen atoms; the scattering power of the oxygen atom is supposed to fall off more rapidly, with increasing glancing angle, than the scattering power of the silicon and aluminium atoms.

Discussion of the Structure.

On account of the unsymmetrical form of the unit cell and the complexity of the structure, it is difficult to show in a diagram the distribution of the silicon and aluminium atoms among the oxygen atoms. In fig. 3 the positions of the silicon and aluminium atoms are projected in a direction parallel to the *a* axis on the (100) plane; fig. 4 represents the appearance of the structure when viewed along the direction of the *b* axis, which is parallel to a cube-edge in the cubic arrangement of oxygens (*cf.* fig. 1).



◍ Silicon }
● Aluminium } The circles indicate the positions, but not the relative sizes of the atoms.

FIG. 3.—Diagram showing Si and Al Atoms projected, in a direction parallel to the a axis, on the (100) plane.

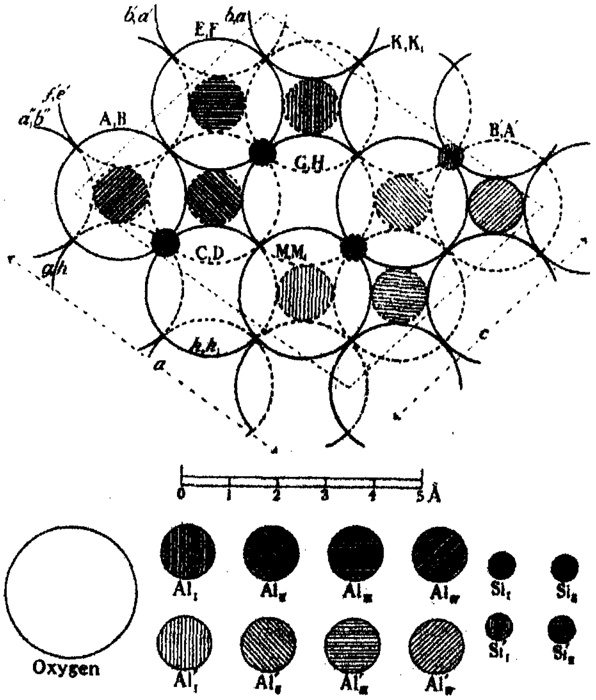


FIG. 4.—Diagram of Structure showing Atoms with Centres lying within the Unit Cell.

EXPLANATION OF FIG 4.

The structure is viewed along the b axis, so that the plane of the diagram is parallel to a cube-face of the underlying oxygen arrangement; the edge of the unit cell parallel to a does not lie in this plane, but its projection, and the edge parallel to c, are shown.

The "ideal" positions of the atoms are shown in this diagram; the displacements from the real positions are small.

All atoms in full line are near the top of the unit cell. Alla toms shown dotted are hidden from direct view by other atoms above them. Oxygen atoms are designated by letters on the circumference of the circles by which they are represented.

Note that each O atom has another O atom lying at a distance $\frac{1}{2}b$ below it; thus B below A, D below C, F below E, H below G; K_1 below K (produced by operation of the symmetry centre midway between K and K_1); M_1 below M (produced by operation of the symmetry centre midway between M and M_1).

A′ denotes an atom obtained from A by the operation of the symmetry centre at the centre of the unit cell; and similarly for B′, C′ ..., $\mathrm{Si}'_{\mathrm{I}}$ ..., $\mathrm{Al}'_{\mathrm{I}}$ .... a denotes an atom obtained from A by the operation of a primitive lattice translation; similarly for b, c, ... and for a', b', c', ... obtained from A′, B′, C′ ....

Si_{I} lies *above* A, k, and *below* oxygen atoms which are at heights $\frac{1}{2}b$ above C, g.

$\mathrm{Si}_{\mathrm{II}}$ lies *below* M, E, and *above* C, G.

In each case the silicon atom lies at the centre of the regular tetrahedron formed by the oxygen atoms specified. The positions of $\mathrm{Si}'_{\mathrm{I}}$ and $\mathrm{Si}'_{\mathrm{II}}$ are obtained by inversion across the symmetry centre lying at the centre of the unit cell.

Al_{I} lies at the centre of six oxygen atoms G F M_1 K_1 b H.

$\mathrm{Al}_{\mathrm{II}}$ lies at the centre of six oxygen atoms C B k_1 M_1 F D.

Note that Al_{I} and $\mathrm{Al}_{\mathrm{II}}$ lie at the same level.

$\mathrm{Al}_{\mathrm{III}}$ lies at the centre of the six oxygen atoms E f' C G b' F.

$\mathrm{Al}_{\mathrm{IV}}$ lies at the centre of the six oxygen atoms B a'' h D e' and the atom $\frac{1}{2}b$ below B.

Note that $\mathrm{Al}_{\mathrm{III}}$ and $\mathrm{Al}_{\mathrm{IV}}$ lie at different levels of the structure.

a'' b'' are obtained from a' b' by a single lattice translation, and so from A′ B′ by two lattice translations.

A simple calculation shows that both silicon and aluminium atoms are distributed in a uniform manner among the interstices of the oxygen arrangement; this was anticipated (it is found in other structures of the type which have been analysed), and might have served as a useful guide in the structure determination had the simple assumptions detailed in a previous section proved inadequate to give a unique solution.

It is also interesting to note that the tetrahedral (SiO_4) groups are distinct, *i.e.*, in no case is an oxygen atom common to two such groups. This configuration has already been found in Mg_2SiO_4,* $ZrSiO_4$,† $Ca_3Al_2Si_2O_{12}$,‡ and the analysis of Be_2SiO_4§ is based on the assumption that the (SiO_4) group occurs as a distinct unit. In crystals where the ratio of the numbers of atoms of silicon and oxygen *is greater* than one to four, neighbouring (SiO_4) groups share one oxygen atom (as in $Be_3Al_2Si_6O_{18}$)[*] or more than one. Apparently, whenever this ratio between the numbers is not greater than one to four, the (SiO_4) groups are distinct. This may serve as a useful guide in the analysis of other structures, *e.g.*, in the attempt which is being made to find the structures of sillimanite and andalusite, the other forms of Al_2SiO_5, which do not appear to be based on a close-packed arrangement of oxygen atoms, and which will therefore present much more complex problems for analysis.

---

* Bragg and Brown., 'Z. f. Krist,' vol. 63, p. 538 (1926).

† Vegard, 'Phil. Mag.,' vol. 7, p. 1151 (1926); Binks, 'Min. Mag.' (1926).

‡ Menzer, 'Z. f. Krist.,' vol. 63, p. 157 (1926).

§ Bragg, 'Roy. Soc. Proc.,' A, vol. 113, p. 642 (1927).

Fig. 5 represents the structure viewed along the *b* axis (as in fig. 4), only the *positions* of the atoms, and not their relative sizes, being indicated. This diagram shows very clearly that there exist, parallel to (100), planes which are densely packed with silicon, aluminium and oxygen atoms, and which are widely spaced. The atoms in one such plane will be tightly bound, but the forces between successive atomic planes will be comparatively weak. This arrangement of all the atoms in widely spaced, densely packed sheets explains the highly perfect cleavage parallel to (100).[†]

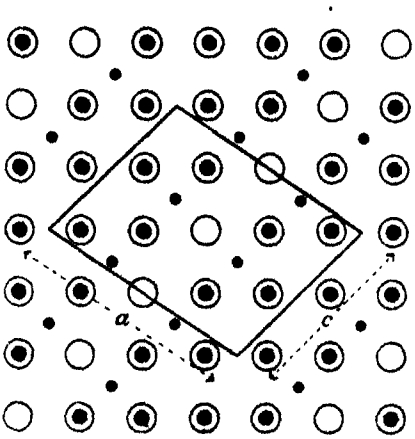


Fig. 5.—Structure viewed along the *b* Axis, showing the Cleavage Planes parallel to (100).

○ Oxygen. ● Aluminium. • Silicon.

Considering only the four silicon atoms and eight aluminium atoms in the unit cell, fig. 3 shows that these are arranged in thin layers, parallel to (010), which are densely packed; successive layers are separated by a spacing much greater than the thickness of each layer. We may then expect a fairly distinct cleavage parallel to (010); this will be less perfect than the (100) cleavage because the sheets parallel to (010) are not so densely packed, and also because the oxygen atoms do not all lie in these atomic sheets. Parallel to (001) no such densely packed layers can be distinguished, so we expect no (001) cleavage to occur.

---

[*] Bragg and West, 'Roy. Soc. Proc.,' A, vol. 109, p. 405 (1925).

[†] A similar configuration is found in the case of the cleavage plane of graphite, *cf.* 'X-rays and Crystal Structure,' p. 231 (5th edition).

These suggestions are in accord with crystallographic data, and also explain why a thin slip of crystal (produced by cleavage parallel to (100)) always splits up into long fine needles parallel to the c axis, and never produces needles parallel to the b axis.

An obvious extension of this argument leads at once to the explanation of the variation of hardness in different directions on the (100) face. A scratching point moving on the (100) face parallel to the c axis easily breaks up the atomic arrangement, owing to the existence of the cleavage parallel to (010), and so the hardness value is small. In a direction parallel to the b axis, however, the hardness value is large since there is no tendency to cleavage parallel to (001). For intermediate directions the hardness will assume intermediate values.

In general, if a crystal face is intersected by cleavage planes, the hardness often has a slightly lower value in a direction parallel to the cleavage planes than in the direction at right angles; these variations are usually much less than the intervals between successive numbers on the usual scale of hardness.* The explanation of the hardness difference in the case of cyanite, in terms of the existence of a cleavage, will apparently apply to all such variations; if the crystal structure suggests an explanation of the existence of these cleavages, that is sufficient to account for the variation of hardness. It will be of interest to examine other structures with a view to finding some other example of a variation comparable in amount with that of cyanite.

Summary.

The crystal structure of cyanite Al_2SiO_5 has been determined by means of a qualitative investigation based on data from X-ray photographs obtained by the rotating crystal method. There are four molecules in the unit cell, which is triclinic, with the following axial lengths and angles :—

$$a = 7{\cdot}09\ \text{Å},\ b = 7{\cdot}72\ \text{Å},\ c = 5{\cdot}56\ \text{Å}\ ;\ \alpha = 90^\circ\ 5\tfrac{1}{4}',\ \beta = 101^\circ\ 2',\ \gamma = 105^\circ\ 44\tfrac{1}{2}'.$$

The structure is based on a slightly distorted cubic close packing of oxygen atoms, the silicon and aluminium atoms being distributed in a uniform manner among the interstices of the oxygen arrangement in such a way that independent tetrahedral SiO_4 groups are formed, while each aluminium atom is at the centre of an octahedral group of six oxygen atoms.

---

* Miers 'Mineralogy,' p. 258.

The existence of densely packed atomic planes, separated by a wide spacing, parallel to (100), explains the existence of the very perfect cleavage parallel to this direction. In a similar way, planes parallel to (010), densely packed with silicon and aluminium atoms, and widely spaced, account for the less perfect but clearly marked (010) cleavage. The variation in hardness on the (100) face is explained as being due to the existence of this cleavage. It is apparently possible to extend this explanation to include the examples, which frequently occur, of hardness variation of much smaller amount.

We are indebted to Prof. W. L. Bragg, F.R.S., for much helpful advice throughout the course of the work. One of us (W.H.T.) is in receipt of a grant from the Board of Scientific and Industrial Research.

The Influence of Vortices upon the Resistance experienced by Solids moving through a Liquid.

By W. G. Bickley, M.Sc., Lecturer in Mathematics at Battersea Polytechnic.

(Communicated by G. I. Taylor, F.R.S.—Received March 16, 1928.)

(1) It is not so long ago that it was generally believed that the " classical " hydrodynamics, as dealing with perfect fluids, was, by reason of the very limitations implied in the term " perfect," incapable of explaining many of the observed facts of fluid motion. The paradox of d'Alembert, that a solid moving through a liquid with constant velocity experienced no resultant force, was in direct contradiction with the observed facts, and, among other things, made the lift on an aeroplane wing as difficult to explain as the drag. The work of Lanchester and Prandtl, however, showed that lift could be explained if there was " circulation " round the aerofoil. Of course, in a truly perfect fluid, this circulation could not be produced—it does need viscosity to originate it—but once produced, the lift follows from the theory appropriate to perfect fluids. It has thus been found possible to explain and calculate lift by means of the classical theory, viscosity only playing a significant part in the close neighbourhood (" grenzschicht ") of the solid. It is proposed to show, in the present

paper, how the presence of vortices in the fluid may cause a force to act on the solid, with a component in the line of motion, and so, at least partially, explain drag. It has long been realised that a body moving through a fluid sets up a train of eddies. The formation of these needs a supply of energy, ultimately dissipated by viscosity, which qualitatively explains the resistance experienced by the solid. It will be shown that the effect of these eddies is not confined to the moment of their birth, but that, so long as they exist, the resultant of the pressure on the solid does not vanish.

This idea is not absolutely new ; it appears in a recent paper by W. Müller.* Müller uses some results due to M. Lagally,† who calculates the resultant force on an immersed solid for a general fluid motion. The result, as far as it concerns vortices, contains their velocities relative to the solid. Despite this, the term $-\frac{1}{2}\rho q^2$ only was used in the pressure equation, although the other term, $\rho\, \partial\phi/\partial t$, must exist on account of the motion. (There is, by Lagally's formulæ, no force without relative motion.) The analysis in the present paper was undertaken partly to supply this omission and partly to check the results of some work upon two-dimensional potential problems in general that it is hoped to publish shortly.

We shall here confine ourselves to two cases of two-dimensional flow past circular cylinders, investigating the resultant force on the cylinder by the full pressure equation

$$p = \rho\left(\frac{\partial\phi}{\partial t} - \frac{1}{2}q^2\right). \tag{1.1}$$

(2) The first example we shall consider is that of a cylinder moving uniformly through a liquid, with two vortices in its wake. Adopting the usual artifice, we shall consider a stationary cylinder in a stream.

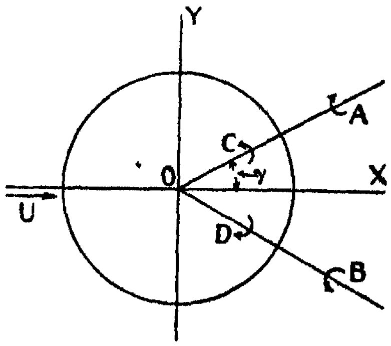


FIG. 1.

If U is the velocity of the stream at infinity, flowing in the positive direction of the x-axis, and there are two symmetrically disposed vortices, of strength κ, and circulations in the senses indicated, at A and B (fig. 1), then to secure no velocity normal to the boundary, we must have images of the vortices at the inverse points, C and D. Denoting the radius of the circle by a, the distance of A and B from the centre by c, and the angle AOX ($=$ BOX) by γ, the velocity potential-stream function $w\ (= \phi + \iota\psi)$, for the two-dimensional motion is given by

---

* 'Z. f. techn. Physik,' vol. 8, p. 62 (1927).

† 'Z. f. angew. Math. u. Mech.,' vol. 2, p. 409 (1922).

$$w = -\mathrm{U}\left(z + \frac{a^2}{z}\right) + \frac{\iota\kappa}{2\pi}\log\frac{\left(z - \frac{a^2}{c}e^{\iota\gamma}\right)\left(z - ce^{-\iota\gamma}\right)}{\left(z - ce^{\iota\gamma}\right)\left(z - \frac{a^2}{c}e^{-\iota\gamma}\right)}, \tag{2.1}$$

z denoting, as usual, $x + \iota y$.

If u and v are the components of the fluid velocity,

$$\begin{aligned} u - \iota v &= -\frac{dw}{dz} \\ &= \mathrm{U}\left(1 - \frac{a^2}{z^2}\right) - \frac{\iota\kappa}{2\pi}\left\{\frac{1}{z - \frac{a^2}{c}e^{\iota\gamma}} - \frac{1}{z - ce^{\iota\gamma}} + \frac{1}{z - ce^{-\iota\gamma}} - \frac{1}{z - \frac{a^2}{c}e^{-\iota\gamma}}\right\}. \end{aligned} \tag{2.21}$$

To find the velocity q on the boundary, put $z = ae^{\iota\theta}$, when

$$\begin{aligned} u - \iota v = \mathrm{U}\left(1 - e^{-2\iota\theta}\right) - \frac{\iota\kappa\left(c^2 - a^2\right)e^{-\iota\theta}}{2\pi a}&\left\{\frac{1}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}\right. \\ &\left. - \frac{1}{c^2 + a^2 - 2ac\cos(\theta + \gamma)}\right\}, \end{aligned}$$

or, taking the conjugate quantity,

$$\begin{aligned} u + \iota v &= \mathrm{U}\left(1 - e^{2\iota\theta}\right) + \frac{\iota\kappa\left(c^2 - a^2\right)e^{\iota\theta}}{2\pi a}\left\{\frac{1}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}\right. \\ &\qquad\left. - \frac{1}{c^2 + a^2 - 2ac\cos(\theta + \gamma)}\right\} \\ &= \iota e^{\iota\theta}\left[-2\mathrm{U}\sin\theta + \frac{\kappa\left(c^2 - a^2\right)}{2\pi a}\left\{\frac{1}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}\right.\right. \\ &\qquad\left.\left. - \frac{1}{c^2 + a^2 - 2ac\cos(\theta + \gamma)}\right\}\right] \end{aligned}$$

or

$$\begin{aligned} q = -2\mathrm{U}\sin\theta + \frac{\kappa\left(c^2 - a^2\right)}{2\pi a}&\left\{\frac{1}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}\right. \\ &\left. - \frac{1}{c^2 + a^2 - 2ac\cos(\theta + \gamma)}\right\}. \end{aligned} \tag{2.22}$$

At this point we will also obtain the velocity of the vortices, which will be required in the sequel. Using subscripts to denote the vortex considered, we can find $u_A - \iota v_A$ by omitting the term $1/(z - ce^{\iota\gamma})$ in (2.21), and putting $z = ce^{\iota\gamma}$ in the remainder. This gives

$$u_A - \iota v_A = U\left(1 - \frac{a^2}{c^2}e^{-2\iota\gamma}\right) + \frac{\iota\kappa}{2\pi}\left\{\frac{ce^{-\iota\gamma}}{c^2 - a^2} - \frac{1}{2\iota c \sin\gamma} - \frac{c}{c^2e^{\iota\gamma} - a^2e^{-\iota\gamma}}\right\}$$

or, taking the conjugate,

$$u_A + \iota v_A = U\left(1 - \frac{a^2}{c^2}e^{2\iota\gamma}\right) + \frac{\iota\kappa}{2\pi}\left\{\frac{ce^{\iota\gamma}}{c^2 - a^2} - \frac{1}{2\iota c \sin\gamma} - \frac{c}{c^2e^{-\iota\gamma} - a^2e^{\iota\gamma}}\right\},$$

whence

$$\left.\begin{aligned} u_A &= U\left(1 - \frac{a^2}{c^2}\cos 2\gamma\right) - \frac{\kappa}{2\pi}\left\{\frac{c\sin\gamma}{c^2 - a^2} + \frac{1}{2c\sin\gamma} - \frac{c(c^2 + a^2)\sin\gamma}{c^4 - 2a^2c^2\cos 2\gamma + a^4}\right\} \\ v_A &= -U\frac{a^2}{c^2}\sin 2\gamma + \frac{\kappa}{2\pi}\left\{\frac{c\cos\gamma}{c^2 - a^2} - \frac{c(c^2 - a^2)\cos\gamma}{c^4 - 2a^2c^2\cos 2\gamma + a^4}\right\} \end{aligned}\right\}. \tag{2.23}$$

Also,

$$u_B = u_A \quad \text{and} \quad v_B = -v_A. \tag{2.24}$$

By the known properties of inversion, the paths of the vortices and those of the images are equally inclined to the radius, so

$$\left.\begin{aligned} u_C &= -\frac{a^2}{c^2}(u_A\cos 2\gamma + v_A\sin 2\gamma) = u_D \\ v_C &= \frac{a^2}{c^2}(v_A\cos 2\gamma + u_A\sin 2\gamma) = -v_D \end{aligned}\right\} \tag{2.25}$$

or

$$\left.\begin{aligned} u_C + \iota v_C &= -(u_A - \iota v_A)\,e^{2\iota\gamma} \\ u_D + \iota v_D &= -(u_A + \iota v_A)\,e^{-2\iota\gamma} \end{aligned}\right\}. \tag{2.26}$$

If the vortices are stationary, *i.e.*, if $u_A = 0 = v_A$, we find, by eliminating U/κ between the two equations of (2.23), and some reduction, that the vortices must lie on the curve whose equation is

$$2c\sin\gamma = c - a^2/c, \tag{2.31}$$

which, interpreted geometrically, means that AC = AB. For any position of the vortex on this curve, the value of κ is given by

$$\kappa = 2Uc\,(1 - a^4/c^4)\sin\gamma. \tag{2.32}$$

(3) The pressure, p, at any point is given by (1.1), and the force on an element $a\,.\,\delta\theta$ of the boundary has components δX and δY, parallel to the x and y axes

respectively, given by $\delta X = -p \,.\, a \,.\, \delta\theta \,.\, \cos\theta$ and $\delta Y = -pa \,.\, \delta\theta \,.\, \sin\theta$. In our case, the component Y vanishes by symmetry, and for X we have

$$X = -a\int_0^{2\pi} p\cos\theta \,.\, d\theta. \qquad (3.1)$$

(*a*) Taking first the q^2 term, we have, from (2.22),

$$q^2 = 4U^2\sin^2\theta$$
$$-\frac{2\kappa U(c^2-a^2)\sin\theta}{\pi a}\left\{\frac{1}{c^2+a^2-2ac\cos(\theta-\gamma)}-\frac{1}{c^2+a^2-2ac\cos(\theta+\gamma)}\right\}$$
$$+\frac{\kappa^2(c^2-a^2)^2}{4\pi^2a^2}\left\{\frac{1}{\{c^2+a^2-2ac\cos(\theta-\gamma)\}^2}+\frac{1}{\{c^2+a^2-2ac\cos(\theta+\gamma)\}^2}\right.$$
$$\left.-\frac{2}{\{c^2+a^2-2ac\cos(\theta-\gamma)\}\,\{c^2+a^2-2ac\cos(\theta-\gamma)\}}\right\}.$$

Using an obvious notation,

$$X_q = \tfrac{1}{2}\rho a\int_0^{2\pi} q^2\cos\theta\, d\theta$$
$$= -2\rho\kappa U\frac{a^2}{c^2}\sin 2\gamma + \frac{\rho\kappa^2}{\pi}\left\{\frac{c\cos\gamma}{c^2-a^2}-\frac{c(c^2-a^2)\cos\gamma}{c^4+a^4-2a^2c^2\cos 2\gamma}\right\}. \qquad (3.21)$$

This agrees with the result obtained by Müller from Lagally's formula. We may note, in comparing it with (2.23), that it may be written

$$X_q = 2\kappa\rho v_A = \kappa\rho\,(v_A - v_B). \qquad (3.22)$$

The general theory referred to above shows that, with any system of vortices, κ_r being the strength, and u_r and v_r the velocity components, of the *r*th vortex,

$$X_q + \iota Y_q = \iota\rho\,\Sigma\kappa_r\,(u_r + \iota v_r). \qquad (3.3)$$

(*b*) Since the above only gives a resultant force when the vortices move, we cannot neglect the $\partial\phi/\partial t$ term. We proceed, therefore, to calculate its contribution to the resultant force.

Assuming the stream velocity, and also the strengths of the vortices, to remain constant (in actual fact, the rate of decay of the eddies is small compared with their motion) we have, rewriting (2.1) with an obvious notation,

$$w = \frac{\iota\kappa}{2\pi}\log\frac{(z-z_C)(z-z_B)}{(z-z_A)(z-z_D)},$$

and therefore

$$\frac{\partial w}{\partial t} = \frac{\iota\kappa}{2\pi}\left\{\frac{u_A+\iota v_A}{z-z_A}-\frac{u_B+\iota v_B}{z-z_B}-\frac{u_C+\iota v_C}{z-z_C}+\frac{u_D+\iota v_D}{z-z_D}\right\}. \qquad (3.41)$$

Using the values of the u's and v's from (2.24) and (2.26), and putting $z = ae^{\iota\theta}$ and inserting the appropriate values of z_A, z_B, z_C, and z_D, we have for the value of $\partial w/\partial t$ on the boundary

$$\frac{\partial w}{\partial t} = \frac{\iota\kappa}{2\pi}\left\{\frac{u_A + \iota v_A}{ae^{\iota\theta} - ce^{\iota\gamma}} + \frac{a(u_A - \iota v_A)e^{2\iota\gamma}}{c(ce^{\iota\theta} - ae^{\iota\gamma})} - \frac{a(u_A + \iota v_A)e^{-2\iota\gamma}}{c(ce^{\iota\theta} - ae^{-\iota\gamma})} - \frac{u_A - \iota v_A}{ae^{\iota\theta} - ce^{-\iota\gamma}}\right\}. \quad (3.42)$$

Picking out the real part,

$$\frac{\partial \phi}{\partial t} = \frac{\kappa}{2\pi c}\left\{\frac{\{2ac\cos\gamma\sin(\theta-\gamma) - (c^2-a^2)\sin\gamma\}u_A + \{2ac\sin\gamma\sin(\theta-\gamma) + (c^2-a^2)\cos\gamma\}v_A}{c^2 + a^2 - 2ac\cos(\theta-\gamma)}\right.$$

$$\left. - \frac{\{2ac\cos\gamma\sin(\theta+\gamma) + (c^2-a^2)\sin\gamma\}u_A + \{2ac\sin\gamma\sin(\theta+\gamma) - (c^2-a^2)\cos\gamma\}v_A}{c^2 + a^2 - 2ac\cos(\theta+\gamma)}\right\}. \quad (3.43)$$

For the resultant force,

$$X_\phi = -\rho a\int_0^{2\pi}\frac{\partial\phi}{\partial t}\cos\theta\,d\theta$$

$$= 2\pi\rho\frac{a^2}{c^2}(u_A\sin 2\gamma - v_A\cos 2\gamma). \quad (3.5)$$

Combining the results from (*a*) and (*b*), we have, finally,

$$X = X_a + X_\phi = 2\kappa\rho\left\{u_A\frac{a^2}{c^2}\sin 2\gamma + v_A\left(1 - \frac{a^2}{c^2}\cos 2\gamma\right)\right\}, \quad (3.61)$$

$$= \frac{\rho a^2\kappa^2}{\pi c^3}\cos\gamma\,\frac{4c^4\sin^2\gamma - (c^2-a^2)^2}{c^4 + a^4 - 2a^2c^2\cos 2\gamma}. \quad (3.62)$$

Comparing this with (2.31), we see that the resultant force is a "drag" if γ is greater than the value for which the vortices are stationary—and that it vanishes when they are stationary.

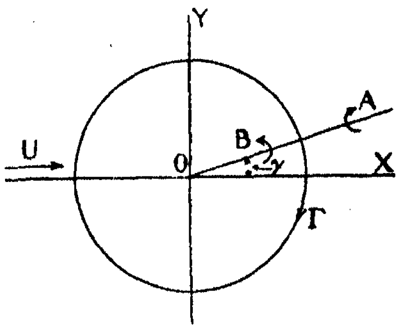


FIG. 2.

(4) The other case we shall consider is when there is one vortex outside the cylinder, and also circulation round it. Here, if κ is the strength of the vortex at A (fig. 2), in the sense indicated, and Γ the circulation (also in the negative direction, so that the lift may come out positive), we must have a vortex at B, the inverse of A, and also a vortex of strength $(\Gamma + \kappa)$

at the centre. The velocity potential-stream function is then (notation as before),

$$w = -\mathrm{U}\left(z + \frac{a^2}{z}\right) - \frac{\iota(\Gamma + \kappa)}{2\pi}\log z - \frac{\iota\kappa}{2\pi}\log\frac{z - ce^{\iota\gamma}}{z - \dfrac{a^2}{e}e^{\iota\gamma}}, \tag{4.1}$$

whence

$$u - \iota v = -\frac{dw}{dz} = \mathrm{U}\left(1 - \frac{a^2}{z^2}\right) + \frac{\iota(\Gamma + \kappa)}{2\pi z}$$

$$+ \frac{\iota\kappa}{2\pi}\left\{\frac{1}{z - ce^{\iota\gamma}} - \frac{1}{z - \dfrac{a^2}{c}e^{\iota\gamma}}\right\}. \tag{4.21}$$

On the boundary, putting $z = ae^{\iota\theta}$,

$$u - \iota v = \mathrm{U}\left(1 - e^{-2\iota\theta}\right) + \frac{\iota(\Gamma + \kappa)}{2\pi a}e^{-\iota\theta} - \frac{\iota\kappa}{2\pi a}\frac{(c^2 - a^2)\,e^{-\iota\theta}}{c^2 + a^2 - 2ac\cos(\theta - \gamma)},$$

or

$$u + \iota v = \mathrm{U}\left(1 - e^{2\iota\theta}\right) - \frac{\iota(\Gamma + \kappa)}{2\pi a}e^{\iota\theta} + \frac{\iota\kappa}{2\pi a}\frac{(c^2 - a^2)\,e^{\iota\theta}}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}$$

$$= \iota e^{\iota\theta}\left\{-2\mathrm{U}\sin\theta - \frac{\Gamma + \kappa}{2\pi a} + \frac{\kappa}{2\pi a}\frac{c^2 - a^2}{c^2 + a^2 - 2ac\cos(\theta - \gamma)}\right\}. \tag{4.22}$$

Also,

$$u_\mathrm{A} - \iota v_\mathrm{A} = \mathrm{U}\left(1 - \frac{a^2}{c^2}e^{-2\iota\gamma}\right) + \frac{\iota(\Gamma + \kappa)}{2\pi c}e^{-\iota\gamma} - \frac{\iota\kappa}{2\pi}\frac{ce^{-\iota\gamma}}{c^2 - a^2},$$

so that

$$u_\mathrm{A} + \iota v_\mathrm{A} = \mathrm{U}\left(1 - \frac{a^2}{c^2}e^{2\iota\gamma}\right) - \frac{\iota(\Gamma + \kappa)}{2\pi c}e^{\iota\gamma} + \frac{\iota\kappa}{2\pi}\frac{ce^{\iota\gamma}}{c^2 - a^2},$$

or

$$\left.\begin{aligned} u_\mathrm{A} &= \mathrm{U}\left(1 - \frac{a^2}{c^2}\cos 2\gamma\right) + \left\{\frac{\Gamma + \kappa}{2\pi c} - \frac{\kappa c}{2\pi(c^2 - a^2)}\right\}\sin\gamma \\ v_\mathrm{A} &= -\mathrm{U}\frac{a^2}{c^2}\sin 2\gamma - \left\{\frac{\Gamma + \kappa}{2\pi c} - \frac{\kappa c}{2\pi(c^2 - a^2)}\right\}\cos\gamma \end{aligned}\right\}, \tag{4.23}$$

and, as before,

$$u_\mathrm{B} + \iota v_\mathrm{B} = -\frac{a^2}{c^2}(u_\mathrm{A} - \iota v_\mathrm{A})\,e^{2\iota\gamma}. \tag{4.24}$$

(5) The resultant force due to the $-\frac{1}{2}\rho q^2$ term has components

$$\mathrm{X} = \tfrac{1}{2}\rho a\int_0^{2\pi} q^2\cos\theta\,.\,d\theta, \qquad \mathrm{Y} = \tfrac{1}{2}\rho a\int_0^{2\pi} q^2\sin\theta\,.\,d\theta. \tag{5.1}$$

Now

$$q^2 = 4U^2 \sin^2\theta + \frac{(\Gamma+\kappa)^2}{4\pi^2a^2} + \frac{\kappa^2}{4\pi^2a^2}\frac{(c^2-a^2)^2}{\{c^2+a^2-2ac\cos(\theta-\gamma)\}^2}$$
$$+\frac{2(\Gamma+\kappa)U}{\pi a}\sin\theta - \frac{2U\kappa}{\pi a}\frac{(c^2-a^2)\sin\theta}{c^2+a^2-2ac\cos(\theta-\gamma)}$$
$$-\frac{\kappa(\Gamma+\kappa)}{2\pi^2a^2}\frac{c^2-a^2}{c^2+a^2-2ac\cos(\theta-\gamma)},$$

and performing the integrations,

$$X_q = \tfrac{1}{2}\rho a\left\{\frac{\kappa^2 c\cos\gamma}{\pi a(c^2-a^2)} - \frac{2U\kappa a}{c^2}\sin 2\gamma - \frac{\kappa(\Gamma+\kappa)}{\pi ac}\cos\gamma\right\}$$
$$= \rho\kappa v_A, \tag{5.21}$$

and

$$Y_q = \tfrac{1}{2}\rho a\left\{\frac{\kappa^2 c\sin\gamma}{\pi a(c^2-a^2)} + \frac{2(\Gamma+\kappa)U}{a}\right.$$
$$\left. - \frac{2\kappa U}{a}\left(1-\frac{a^2}{c^2}\cos 2\gamma\right) - \frac{\kappa(\Gamma+\kappa)}{\pi ac}\sin\gamma\right\}$$
$$= \rho(\Gamma+\kappa)U - \rho\kappa u_A. \tag{5.22}$$

Combining (5.21) and (5.22) we derive

$$X_q + \iota Y_q = \iota\rho\{(\Gamma+\kappa)U - \kappa(u_A + \iota v_A)\}. \tag{5.23}$$

We see here the result already given in (3.3), together with another term which contains the product of the stream velocity and the circulation at infinity—a term also given by the general methods already referred to in connection with (3.3)—which is, of course, the mathematical expression of the Lanchester-Prandtl theory of lift, and of the so-called Magnus effect.

For the $\partial\phi/\partial t$ term, again considering only the motion of the vortices, we have, picking out the relevant portion of (3.43),

$$\frac{\partial\phi}{\partial t} = \frac{\kappa}{2\pi c}\left\{\frac{\{2ac\cos\gamma\sin(\theta-\gamma)-(c^2-a^2)\sin\gamma\}u_A + \{2ac\sin\gamma\sin(\theta-\gamma)+(c^2-a^2)\cos\gamma\}v_A}{c^2+a^2-2ac\cos(\theta-\gamma)}\right\} \tag{5.4}$$

whence

$$X_\phi = -\rho a\int_0^{2\pi}\frac{\partial\phi}{\partial t}\cos\theta\,d\theta = \rho\kappa\frac{a^2}{c^2}(u_A\sin 2\gamma - v_A\cos 2\gamma) \tag{5.51}$$

and

$$Y_\phi = -\rho a\int_0^{2\pi}\frac{\partial\phi}{\partial t}\sin\theta\,d\theta = \rho\kappa\frac{a^2}{c^2}(u_A\cos 2\gamma + v_A\sin 2\gamma) \tag{5.52}$$

so that

$$X_\phi + \iota Y_\phi = -\iota\rho\kappa\frac{a^2}{c^2}(u_A - \iota v_A)e^{2\iota\gamma}. \tag{5.53}$$

Combining these results we have

$$X + \iota Y = \iota\rho\left\{(\Gamma+\kappa)U - \kappa u_A\left(1+\frac{a^2}{c^2}e^{2\iota\gamma}\right) - \iota\kappa v_A\left(1-\frac{a^2}{c^2}e^{2\iota\gamma}\right)\right\}. \tag{5.61}$$

In the language of aerodynamics,

$$\begin{aligned} X = \text{Drag} &= \rho\kappa\left\{u_A\frac{a^2}{c^2}\sin 2\gamma + v_A\left(1-\frac{a^2}{c^2}\cos 2\gamma\right)\right\} \\ &= \rho\kappa\left(1-\frac{a^2}{c^2}\right)\left\{\frac{\kappa c}{2\pi(c^2-a^2)} - \frac{(\Gamma+\kappa)}{2\pi c}\right\}\cos\gamma \end{aligned} \tag{5.62}$$

$$\begin{aligned} Y = \text{Lift} &= \rho\left\{(\Gamma+\kappa)U - \kappa u_A\left(1+\frac{a^2}{c^2}\cos 2\gamma\right) - \kappa v_A\frac{a^2}{c^2}\sin 2\gamma\right\} \\ &= \rho\left[(\Gamma+\kappa)U - \kappa\left\{U\left(1-\frac{a^4}{c^4}\right) - \left(\frac{\kappa c}{2\pi(c^2-a^2)} - \frac{\Gamma+\kappa}{2\pi c}\right)\sin\gamma\right\}\right] \end{aligned} \tag{5.63}$$

(6) If, in the above, we put $(\Gamma + \kappa) = 0$, making the circulation round the boundary equal and opposite to that of the vortex (and consequently the circulation round any simple circuit embracing both cylinder and vortex zero), and also put $\gamma = 0$, we find

$$\text{Drag} = \frac{\rho\Gamma^2}{2\pi c} \quad \text{and} \quad \text{Lift} = \rho\Gamma U\left(1-\frac{a^4}{c^4}\right). \tag{6.1}$$

Introducing the coefficients of drag and lift, defined by $C_D = \text{Drag}/\rho S U^2$, $C_L = \text{Lift}/\rho S U^2$, S being the projected area, we find

$$C_D = \frac{1}{4\pi}\frac{a}{c}\left(\frac{\Gamma}{aU}\right)^2, \qquad C_L = \tfrac{1}{2}\left(1-\frac{a^4}{c^4}\right)\frac{\Gamma}{aU} \tag{6.2}$$

The lift-drag curves plotted from these are a family of parabolas. In fig. 3 are given three published curves for Flettner rotors (full lines) and curves calculated from (6.2) for the values $2a$, $4a$, and $6a$ of c. The agreement is seen to be fair, except for low lifts.

More generally, keeping $(\Gamma + \kappa) = 0$, but retaining γ, we find

$$C_D = \frac{\cos\gamma}{4\pi}\frac{a}{c}\left(\frac{\Gamma}{aU}\right)^2, \tag{6.31}$$

$$C_L = \tfrac{1}{2}\left(1-\frac{a^4}{c^4}\right)\frac{\Gamma}{aU} - \frac{\sin\gamma}{4\pi}\frac{a}{c}\left(\frac{\Gamma}{aU}\right)^2. \tag{6.32}$$

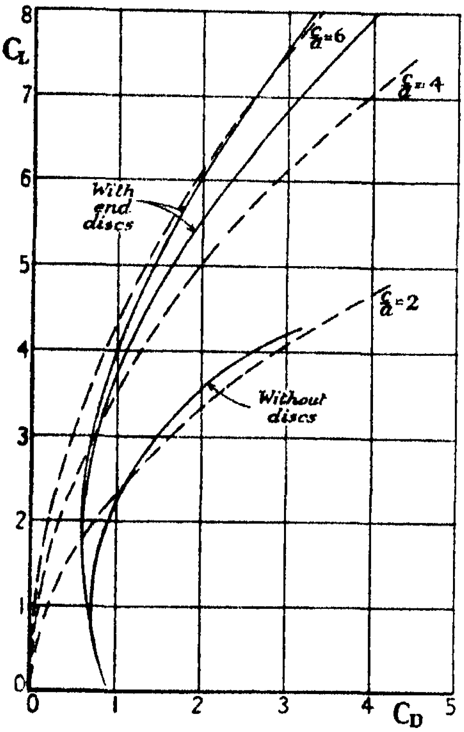


FIG. 3.

The lift-drag curves obtained from these equations are all parabolas. In fig. 4 are given curves for $c = 2a$, for values $-30°$, $-15°$, $0°$, $15°$ and $30°$ of γ; and in fig. 5 the corresponding curves for $c = 6a$. It is seen that the shape of these for positive values of γ is much closer to that of the experimental curves of fig. 3. There still remains an approximately constant difference between the drag coefficient of the experimental curve and the calculated curve agreeing most closely with it in shape. This was to be expected, since we have taken no account of the conditions in the "grenzschicht," where the true viscous resistance plays a predominant part, and where the vortices originate. Moreover, in practice, a train of vortices is thrown off, and it is not to be expected that a single vortex can entirely simulate their effect. It is, however, significant that the curves for the rotor not fitted with discs should be best represented by a small value of c/a, while those for the rotors with discs require a larger value of this ratio. In the former case the motion is not approximately two-dimensional—trailing vortices, roughly perpendicular to the axis of the cylinder, are thrown off from the ends of the cylinder. The portions of these in the close

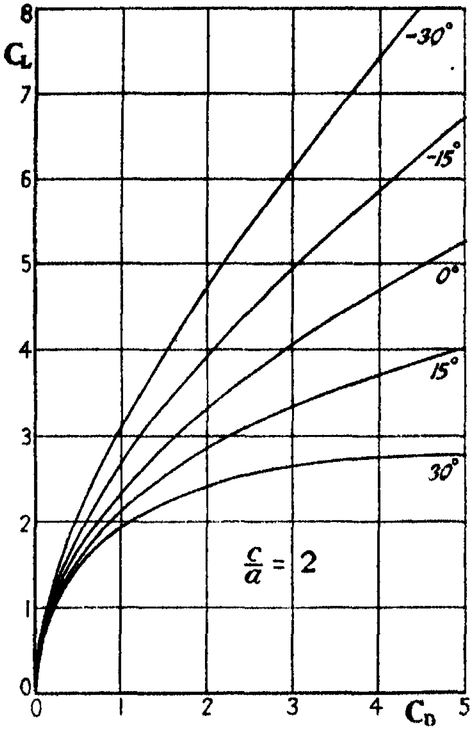


FIG. 4.

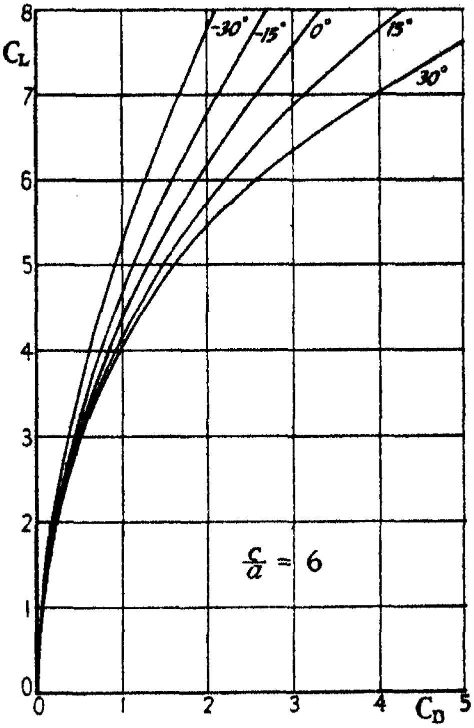


FIG. 5.

neighbourhood of the rotor will have considerable effect, and it would appear probable, therefore, that a small value of c/a would be required to represent it. The effect of vortices considered here is not, however, quite the same thing as that leading to the " induced drag " of aerodynamics.

In conclusion, I wish to thank Mr. H. H. Taylour for his help in preparing the diagrams.

Summary.

Two cases of two-dimensional streaming past a cylinder, with vortices present, are worked out in detail, and the resultant of the fluid pressures calculated. In the first, where there are two symmetrically disposed vortices, the resultant is a force in the line of motion. In the second, with one vortex and circulation, the force is inclined to the direction of motion. " Lift " and " drag " are calculated, and the lift-drag curves show a resemblance to those found experimentally for the Flettner rotor.

The Origin of Speech—A Hypothesis.

By Sir Richard A. S. Paget, Bart.

(Communicated by Sir William Bragg, F.R.S.—Received June 13, 1927.)

In a previous paper* experiments were described in which vowel-like sounds were produced by models consisting of two suitably tuned Helmholtz Resonators, connected in series, and energised by a reed of variable pitch. It was found that with these models, as also in some whose resonators were connected in parallel, the vowel-like sounds were, in general, but little affected by variation of pitch of the energising reeds, though it was evident that the resultant sound, if analysed, would necessarily be very different when the relations of the energising reed-note to the resonant notes of the two cavities was changed.

In a subsequent paper,† observations were described which indicated that in recognising the consonant sounds k/g and t/d, the human ear is not directly interested in the resonant changes, which constitute these sounds, and which vary very greatly according to the upper resonance of the vowel with which the consonant is associated. The ear appears rather to be on the look out for acoustic indications of the gestures of the organs of articulation, by which these consonants are produced.

The conclusion was drawn that human speech was essentially a branch of human gesture which the ear had learnt to identify by means of the secondary effects of the gestures in modifying the resonances of the gesticulating cavities.

It has been observed by previous writers‡ that, in many different languages, words meaning "little" or "near" contained the vowel sound i (as in it), while words meaning "big" or "far" contained the vowels ɑ, o or u. These observations were explainable on the pantomimic gesture theory by the fact that the vowel i is due to the formation of a *small* front resonator—made between the front of the tongue and the lips—or (in other words) to a *near* position of the tongue to the lips, while the ɑ, o and u sounds are due to a *large* front resonator or to a *far* position of the tongue from the lips.

The suggestion was therefore hazarded that the original word in human speech for little was i–i (ee–ee), while that for big was ɑ–ɑ or ɔ–ɔ (aw–aw). In reference to this suggestion, Dr. Neville Whymant recently pointed out to the present writer that, in primitive Polynesian and archaic Japanese, the words for little and big were respectively I′ I and ŌHŌ. In view of this apparent confirmation, the attempt was made to devise a number of fabricated words, descriptive of various simple actions, such as primitive man might have used. The process adopted was to make, in the first instance, a pantomimic hand gesture, descriptive of the action in question, and then to invent a corresponding gesture of the tongue and lips which might do duty in the place of the hand gesture.* When a satisfactory tongue and lip gesture had been evolved, it was then converted into whispered speech, by blowing air through the mouth, or into voiced speech by humming or grunting during the performance of the gesture of the tongue and lips, etc.

---

* 'Roy. Soc. Proc.,' A, vol. 102, p. 752 (1923).

† 'Roy. Soc. Proc.,' A, vol. 106, p. 150 (1924).

‡ E. B. Tylor, 'Primitive Culture,' vol. 1, p. 220. Prof. Otto Jespersen, 'Language: its Nature, Development and Origin,' p. 402 (1922).

In the following table the fabricated words thus formed are set out, together with Dr. Neville Whymant's additions, in the 4th column, of some actual words with similar meaning, occurring in primitive Polynesian, Melanesian, Indonesian or archaic Japanese speech.

Action to be Described.	Pantomimic Gesture of Tongue.	Phonetic Resultant.	Polynesian, Melanesian, Indonesian or archaic Japanese words of similar significance, supplied by Dr. Neville Whymant.
Reach up	Tongue reaching up to touch the palate	ɑd, ɑð or ɑθ	ɑdɑ, ɑdhɑ.—*ɑdɑru, idɑru*.
Feel, Stroke	Tongue feeling palate backwards and downwards	θrɑ, ðrɑ, ırɑ, drɑ	tura, tulɑ.—*tɑtɑru*.
Feel, Smooth or Stroke up	Tongue feeling behind lower teeth and up behind front teeth	ɑil, oil, əil	ɑilɑ, ϵilɑ.—*tɑirɑ*.
Draw back suddenly	Tongue protruded and withdrawn	ærʌp, æðʌp	ϵðupu, ðupi.
Scrape	Tongue scraped between the teeth	ðʌp, ðup, ðu	ðubu, ðuu.
Wave Aloft	Tongue waved, touching palate	leðl–leðl or ledl–ledl	lete–lete, vele-vele.
Shake (like a mat)	Tongue shaken behind teeth	ɒlɑ–ɒlɑ	ore–ore.
Stab or Spear	Tongue protruded between lips and teeth	peð, pʌð or pʌl	bulu, mbulu.
Shoot (bow and arrow)	Tongue reflexed, grip at back of tongue and suddenly release	ðr-ki or ðr-kui or dr-ki, dr-kui	no similar word discovered.
Pull Down	Tongue reflexed and lowered	trɑ, trɔ	terei, ndrei—*tore, ɑtɑru*.

---

* Dr. Gregory (Met. Mus. Nat. Hist., New York) pointed out that the original gesture was more likely to have been a generalised one (including a mouth gesture) and that the concomitant bodily gestures were gradually dropped.

In the column labelled Phonetic Resultant, the sounds produced by the pantomimic gestures are transcribed as originally recorded.

There are, however, a number of additional resultants, due to the same tongue and lip gestures, which might equally have been included, and which, in some cases, give an even better fit to Dr. Whymant's words in the fourth column.

Thus D, T and N are all due to the same tongue gesture—the differentiations being made farther back. Therefore, " dra " for " feel, stroke," might equally have been written " tra "—which closely resembles Dr. Whymant's " tura."

Similarly, P, B and M are due to a single lip gesture, so that ðʌp, ðup might also be ðʌb, ðub—a closer fit to Dr. Whymant's ðubu.

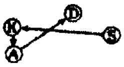
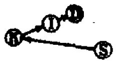
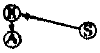
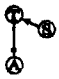
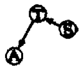
Ledl–ledl might equally be letl–letl, so as to correspond with lete–lete, " to wave aloft," while pʌl written as bʌl or mbʌl, compares very closely with the actual words bulu and mbulu.

It will be seen that in every instance except one, Dr. Whymant has found actual words of similar meaning which show a close phonetic relation to the fabricated words. For the present purpose, only the closest " fits " given by Dr. Whymant have been included in the table.

In view of these results and at the suggestion of Professor L. H. Gray of Columbia University, an examination was made of some Aryan Roots. Taking the list of these given in Professor Walter Skeat's Etymological Dictionary of the English Language, the following were selected as giving indication of the same principle of pantomimic gesture.

Root.	Meaning.	As in :	Pantomimic Gesture.
AGH or ANGH	**choke, strangle**	anguish	**back of tongue tightened against back of throat.**
AD	**eat**	*edo* (Lat.)	*tongue touching palate as in eating.*
AN	**breathe**	animal	**breathing in through mouth and out through nose.**
AP, AM	**seize, take**	redeem	**seizing gesture of jaws and tongue or lips.**
KAR	**curve, roll**	cycle	**curving or rolling (backwards) gesture of the tongue.**
KU	**swell out, be hollow**	cave	**tongue withdrawn to make large cavity, till back of tongue closes against back of throat; followed by plosive release to emphasise resonance of the hollow cavity.**
GAR	**devour, swallow, eat or drink greedily**	**gargle**	**a swallowing or gulping gesture.**
GUS	**choose, taste**	**gust**	**a swallowing and tasting gesture.**
GHAR	**seize, grasp, hold, contain, bend**	**girth**	**grip of back of tongue against soft palate and back of throat, followed by a bending back or enveloping gesture of the tongue.**

Root.	Meaning.	As in :	Pantomimic Gesture.
GHU	pour	gush	mouth formed like a bottle and resonance emphasised by GH release.
TANK	compact, compress	thong	tongue compressed first against front then against back of palate, followed by plosive release.
DA	give	datum	offering gesture made with the tip of the tongue.
DUG	to milk, yield milk	daughter	sucking, followed by swallowing gesture of the tongue.
PA	to feed, nourish	papa	sucking or smacking gesture made by the lips.
PAK	bind, fasten, fix, hold fast	pack	release at lips, followed by grip at back of tongue and plosive release.
PAP, PAMP	swell out, grow round	pimple	pantomimic inflation of cheeks, followed by release and re-inflation.
BHRAGH	break	break	grip with lips (B), bending back the tongue (HR), dropping tongue (A), and finally tightening and lifting back of tongue against back of throat.
MAD	chew	meat	an eating gesture.
MARG	rub gently, wipe, stroke, milk	milk	stroking gesture—probably beginning originally with tongue between lips and bent back so as to stroke hard palate. Stroking movement is continued back till G closure is produced.
MU	bind, close, shut	mute	oral cavity is formed as an enclosure with closed mouth which is finally opened to the smallest extent to produce the U.
RIGH, LIGH	lick	lick	RI and LI are due to licking gestures. Terminal GH is due to continuing the gesture down and back behind the lower teeth.
RUG, LUG	break, bend	lug	RU and LU are due to bending back tongue. Terminal G is due to raising back of tongue to form a grip against the back of throat. Full gesture resembles that of breaking a stick between the two hands.
LAGH	lie down	low	tongue raised to the palate for L, is suddenly flattened in the mouth for AGH.
WA, AW (originally AWE)	breathe, blow	weather	blowing gesture.
UD, WAD	well, gush out, moisten, wet	water	drinking gesture, as in drinking from the palm of the hand.
WAR, WAL	cover, surround, protect, guard	revolve	lip hold (WA), followed by a protecting gesture of tongue (R or L).
WARG	press, urge, shut in, bind	verge	similar to WAR but followed by a final grip G.
WIK	bind, fasten	vetch	similar to WARG but simpler, the tongue rising to the I position and thence receding to the K closure.
SA	sow, strew, scatter	sow	similar to the giving gesture DA, but substituting the small cavity S (made between the back of the front teeth, the tip of the tongue and roof of mouth) for release D. The contents of the S cavity are scattered by the gesture to the A position.

Root.	Meaning.	As in :	Pantomimic Gesture.
SAR, SAL	keep, preserve, make safe, keep whole or sound	salute	compare WAR, WAL, substituting the initial tongue hold S for the lip hold W.
SIW, SU	sew, stitch together	sew	S hold, followed by withdrawal of tongue to I and protrusion of lips to W or by direct protrusion of lips to U. The S appears to be the grip and I W or U the needle gesture.
SUK	to flow, cause to flow, suck	suck	a sucking and swallowing gesture.
SKAN	cut, dig	canal	the cutting or cleaving gesture SKA is followed by the throwing up gesture N.
SKAG	shake	shake	the hold S is followed by the re-duplicated grip K, G.
SKAD SKID	cleave, scatter	scatter	S hold followed by the rapid movements to K, A or I, and D positions—the points of closure or constriction following a  or  track, which may be significant.
SKA SKAP	cut, hew, chop	chop	the  track forms a good chopping symbol. The significance of the P closure is not obvious.
SKAP	dig	ship	compare SKAN—substituting the upward throw of the jaw (producing a P closure) for that of the tongue to produce N.
STA	stand	station	the tongue track is of the type  or  which may denote a standing position.
STARG STRAG STRIG STRUG	stretch	stretch	ST may be the hold, R the backwards stretching movement, and G the hold at the other end.
STIGH	stride, climb, ascend	stile	the tongue ascends to the T and I and falls back to the low GH position.
SNA	bind together, fasten	sinew	the tongue-hold S is drawn back tight at N and falls back to A. Compare the gestures of SIW, SU.
SNAR SNARK	twist, draw tight, entwine, make a noose	snare	The SNA gesture is completed by the bending back gesture R with or without the final hold K.
SPA	draw out, extend, increase	spin	the small cavity S, closed at P, is suddenly enlarged to A.
SPU, SPIW	spit out	spue	the pantomimic gesture is obvious.
SMI	smile	smile	the pursed-up mouth SM is broadened into the smiling attitude which produces an I to E vowel.
SRU STRU	flow, stream	stream	the tongue slides along the palate to the U position (the lips being unmoved).

Root.	Meaning.	As in :	Pantomimic Gesture.
SWAP	sleep	soporific	here, conceivably, the small mouth opening SW may represent the human eye, afterwards closed at P. The u vowel (as in Greek ὕπνος would then seem more appropriate than the larger opening of A. Compare BHAS, shine—in which the SWAP gestures are reversed.
SWAR SAR	string, bind	series	S or SW is the original hold. R is the stretching or binding movement.
SWAL	toss, agitate, swell	swell	S is again the initial hold or the initial (small) size, A is the enlarging gesture, L is the tossing or agitating gesture.

While admitting that some of the pantomimic derivations proposed may be fanciful, it is suggested that the words selected, when taken as a whole, do give evidence of the existence of a pantomimic principle of formation.

Taking the consonants in the order of their positions of closure (full or partial), from the front backwards, the following conclusions are suggested :—

P.B. M. BH. These commonly signify closing, containing, gripping or the reverse actions, swell, burst ;

e.g., ***RA*** has the meaning rest (a reclining gesture of the tongue). With an added P or BH the meaning is cover, roof over ; or it may mean snatch, seize—the jaw gesture being then the significant part.

WA means breathe : ***WAM*** means spit out, vomit.

TAR means attain : ***TARP*** is to be satiated.

KAR, curve, roll up—with added M means to be tired—suggesting rolled up in an enclosed space.

PAP, ***PAMP***, swell out, grow—round, are obviously pantomimic.

NAM, count, portion out, suggests the gesture NA or DA, to give, with a portion of the "gift" cut off at M. Compare TAM or TAN to cut, and DAM, TAM, to tame, *i.e.*, put down and enclose.

MA, measure, is a compass-action of the jaws, imitating the action of separating the two hands in measuring.

U. IW. I. Y. The vowel *U* is a projecting, pointing, directing, spouting gesture,

e.g., ***US***, burn.

TU. ***THU***, swell.

KU, swell.

DU. ***DUK***. ***TUH***, lead, conduct.

DU. FU, stink, be foul (blowing out gesture); or purify, cleanse; or beget (ejecting gesture); compare SU, generate, produce.

GU, bellow, pour.

GHUS, gush.

PAU. FAU, little—probably to make little, *i.e.*, by reduction of the A to U.

MU, move, push.

BHU, grow.

LU, cut off, separate, loosen.

IW as in *DIW, TIW*, shine, may possibly be a lip gesture, imitating the human eye; compare *NIK*, let fall, wink, where the fall of the eyelid is generalised; also *WID*, see, observe—all of these are made with a *miniature* lip opening.

WA. AW, breath and *WAD, UD*, to well or gush out, seem to be primarily in-breathing or in-drawing gestures, which have been converted into speech sounds by reversing the air flow.

WI, go, drive, is a lip-pointing gesture; so is *YA*, that, and *YA*, go.

RI, LI, pour, distil, melt, flow, are small cavity gestures like *MI*, diminish.

S is due to a gesture of forward constriction and represents generally a holding (forward) or bringing to a point or jet.

Compare the following :—

AR	move	ARS	flow.
YA	go	YAS	ferment, seethe or YAS gird.
LA	be low	LAS	pick out, glean (*i.e.*, bend low and hold in front).
LU	cut off	LUS	to be loose (*i.e.*, held, cut off).
GA	come	GAS	bring, heap together.
GHAR	rub................	GHARS	bristle (*i.e.*, be brought to a point S).
GHU	pour	GHUS	gush.
DHWAR	rush forward........	DHWAS	Fall, perish (*i.e.*, rush forward to "constriction")
PAL	cover	SPAL	quiver, jerk, struggle.
PRU	spring up	PRUS	burn (*i.e.*, spring up in jets of flame).

MU	move, push, strip off	MUS	steal (*i.e.*, strip off and hold)
MARD	rub down	SMARD	pain, cause to smart.
KAR	be hard, rough	SKAR	cut, scrape, furrow.
KAR	move, curve, roll	SKAR	move hither and thither (*i.e.*, being held).
KUBH	going down	SKUBH	become agitated.
KU	swell	SKU	cover, shelter (*i.e.*, by hand)
KRU	be hard	SKLU	shut (*i.e.*, hold hard).
WA	breathe, blow	SWA	drive, toss, SWAL, agitate.
WAD	well, gush out	SWAD	please, be sweet (? holding that which wells out).
WAR WAL	cover, surround, protect	SWAR	string, bind.
		SWAL	toss, agitate, swell.
TA	stretch	STA	stand (*i.e.*, hold stretched).
TAK	fit, prepare, make	STAK	stick, stand fast (*i.e.*, with added grip).
TAR	pass over or through	STAR	strew, spread out (compare KU, swell and SKU, cover).
TARGH	pull	STARG	stretch (compare STAR, strew, spread out + G. *q.v.*).
TU	swell	STU	make firm (*i.e.*, hold swelled).
		STU	strike (grip S made to swell).

SA	sow, scatter (S, grip, followed by tongue movement to A).
SUS	to dry, wither (possibly a squeezing dry).
SIW	sew—(hold S, and needle gesture, I–W).

Other words showing the significance of S are US, burn (jets of flame), PIS, pound, and GHAIS, stick, adhere.

TH. DH. These show a smearing, kneading or emollient signification.
e.g. ***TWAK***, ***THWAH***, dip, wash.
DHIGH, smear, knead, mould, form.
DHUGH, milk.

Compare *MAR*, rub, and *MARDH*, be soft, moist, wet.

T. D. N. Initial T and D do not appear to be generally significant. In TA or TAN, stretch, DAR, TAR, tear, rend, the gesture is evident. D and N often relate to eating, tasting, etc., *e.g.*, MAD, chew, NAD, NUD, enjoy, PAT, FAT, feed, BHID, cleave, bite.

Terminal T and D and N commonly denote a closure or hold.

Compare :—

MAR	rub................	MARD	rub down, crush.
MA	(emphasis on M), diminish	MAN	pound, knead, macerate.
MA	measure............	MAT	turn, whirl, throw, spin.
SA	sow, scatter	SAT	full.
SKU	cover, shelter	SKUT SKUT	move, shake, fly, fall, drop. spring out.
RA	go, drive	RAT	split, gnaw, scratch.
GA GHA	gape	GAT GHAD	seize.
KAR	curve, roll..........	KAT	cover, protect.

R, DR, TR. The consonant R gives evidence of having been originally produced by a reflexed movement of the tongue—as in the Wessex R, now common in the United States of America—since it is very generally associated with the meanings, curve, bend back, enclose, bind, etc. DR, TR, frequently denote running, flowing or walking—the direction of motion of the tongue gesture being towards the speaker. The backward bending gesture is found in such words as :—

RA, **rest.**
RAG, **stretch (with G grip).**
ARK, **protect.**
RAP, **cover, roof over.**
RUP, **break, tear, seize.**
KARM, **be tired.**
KRI, **cling to, lean against, incline.**

GAR, grind, crumble, oppress; seize, grasp, contain; bend.
TARK, twist.
DAR, *TAR*, tear, rend.
DAR, sleep, may signify to drop down (DA) and curl up (R).
PRAK, plait, weave, fold together.
BHRAK, shut in, stop up.
BHRAG, break.
MAR, grind.
BHAR, bear, carry.
WARP, throw—where WAR is the backward movement (*cf.* WAR, cover, surround, protect) and P the projection.
SRU, *STRU*, flow.

Compare also the following :—

PA	feed	PAR (PAL)	fill.
MA	(accent on M) diminish	MAR	grind, rub.
TA	stretch	TAR	tear, rend.
DHA	suck	DHAR DHARGH	support, sustain, hold, keep. bear, maintain, retain; also make firm, fasten.
WA	breathe	WAR	speak; also WAR, wind, turn, roll; also drag, tear.
SA	sow, scatter	SAR SWAR	string, bind; also keep, preserve.

Of DR, TR, the following may be instanced :—

DRA, run.
DRAP, flow.
TRAP, tramp.
TRAD, tread.

L. This is due (as is well known) to a tongue gesture very similar to that of R—the difference being that in L the tongue makes more of a point-contact with the palate, so as not to produce the small cavity behind the point of the tongue, as in the reflexed R. L can also be made with protruded tongue touching the front of the upper lip. It is probably that in primitive speech the tongue was, in fact, often so protruded. Thus LUBH (love) would appear to be a phallic tongue gesture, of which the counterpart is KA or KAM (love).

PRI (love) made with protruded tongue, entirely surrounded by the lips, so as to make the plosive P on withdrawal, is probably a similar gesture to LUBH.

Generally speaking, L denotes movement, flow, or rapid change of posture, *e g.*,

LI, pour, distil, melt, flow.
LU, wash.
BHLAGH, strike, beat.
LIGH, lick.
PLAK, *PLAG*, strike.
TAL, lift.
KAL (*KAR*), project, stand up; also hide, cover.
LAK, bend, depress.
LAGH, lie down.
LAD, let go.

Compare :—

TA	stretch	TAL	lift.
PA	feed	PAL	fill.
BHA	shine	BHAL	resound.
BHA	speak clearly	BHLA	blow, puff, spout forth.
BHAG	portion out	BHLAGH	strike, beat.
WA	breathe, blow	WAL	be warm, hot, boil.
SKA	cover, shade, hide ..	SKAL	shear, cut, cleave, scratch, dig (*i.e.*, cover and throw up).
SKU	cover, shelter	SKLU	shut.
SPA	draw out	SPAL	quiver, jerk, struggle, kick, fling, flutter.
SWA	drive, toss..........	SWAL	toss, agitate, swell.

K, *G*, *NK*, *NG*, *GH*, *H*. All of these commonly denote a gripping or constricting action. GH is sometimes associated with the idea of moisture (as in the action of "clearing the throat"). Typical words are :—

KAK, *KANK*, *HANG*, hang.
TANK, *THANG*, contract.
GAR, seize, grasp, contain.
TARGH, pull, draw violently.
DAK, *TAH*, *TANG*, take, hold.
MARG, rub gently, milk.
MIGH, sprinkle, wet (salivary gesture).
RUG, *LUG*, break, bend.

Compare also the following :—

PA	feed	PAK	bind, fasten, fix, hold fast.
BHAR	bear, carry	BHARK BHRAK	shut in, stop up, cram.
MA	think	MAK	have power, be great (*i.e.*, think and hold).
MA	diminish (emphasis on M)	MAK	pound, knead, macerate.
MI	diminish	MIK	mix (*i.e.*, in small cavity).
RA	rest (tongue reclining)	RAG	stretch.
RA	go (tongue moving)..		
RI	pour	RIK	scratch, furrow, tear.
LA	to be low	LAK	bend, depress (*i.e.*, hold down).
		LAG	collect.
		LAGH	lie down (and stay down !).
WAR	cover	WARG	press, urge, shut in, bend.
		WARGH	choke, strangle.
WI	go	WIK	bind, fasten.

More than 200 words in Dr. Skeat's list appear to be pantomimic, and, as will have been seen from the comparisons already made, the gestures are largely systematic. The pantomimic root words seem, in fact, to be built up much in the same way as the Chinese ideographs were formed, viz., by addition of significant elements.

The following are a few additional fabricated words which have not yet been put to the test of comparison with those of primitive speech :—

Action.	Pantomimic Gesture.	Phonetic Resultant.
Sew	a grip with protruded lips—retracted and released (PUI), followed by a new grip at a higher level (S), retracted (R) and brought forward to near the original starting point.	PUI–SRI B ZL M
	Alternative (to second part of gesture) S grip followed by lifting and protruding the tongue and finally curving it downward and back into the mouth.	PUI–SLA B ZR M
Coil or Roll (towards oneself)	initial back grip K followed by further depression of tongue (ɔ), lifting to near palate (I) (and bending back so as to touch the palate before dropping to A.	KɔILA (KAWILA) or KɔIRA or GɔIRA (compare "coil").
Roll (away)	tongue starts from the K, TL, or R position and curves to I, down to ɔ and up from the back to I, before repeating the cycle.	TLIɔe–TLIɔe DR æ–DR æ K K R R

Action.	Pantomimic Gesture.	Phonetic Resultant.
Blow (through a tube)	a sudden (plosive) blast with lips protruded	KY (as in French cul) ŋ (*ng*)
Dance up and down, to and fro	dancing action of the tongue	LI–LƆ R R
Plough	tongue tip drawn back and scraped forward along the floor of the mouth	KƆRI (R guttural) KƆGI KƆI
Strip grains of cereal from the stalk—	"strip" tongue between teeth	θP or ðB.
collect—	reflexed tongue	R
take a handful and scatter it	make a front grip between tongue, teeth and lips and suddenly lower and protrude the tongue.	sɑ, sə or zɑ, zə. The resultant word is therefore :— θ′PRSɑ or ð′BRZɑ, say θʌPəRSɑ or θʌPəRSə or ðʌBeRZɑ or ðʌBeRZə
Pick berries,	front grip closed	SP or SB
collect them	reflexed tongue	R
and bury in the ground	tongue tip buried between front of lower teeth and under lip	DVð or DVZ or TVθ or TVS The resultant word is then :— SB′RDVð or SB′RDVZ or SP′RTVθ or SP′RTVS say SBəRDʌVZ or SPəRTʌVS.

From the last two examples it may be seen that—as linguistic study has already indicated—primitive language is likely to have been very condensed, and that each "word" was probably built up of a succession of significant gestures, equal in meaning to a modern sentence.

It will be also seen how roots—like the Aryan roots—may have arisen ; as for example the fabricated root *pərs*, *bərz*, or *berd*, *pert* (in the two last examples) with the meaning "collect for planting." This part of the word might be expected to remain constant though other parts might change. For example, to pick berries off a high bush, might be "is" (ees)—with high pitched tongue—instead of "s" ; to pick them off the ground might be "las" ; and so on.

This principle of word-building seems to have been recognised by Dr. J. Rae, of Honolulu—writing in 'The Polynesian,' of September 27th, October 4th and October 11th, 1862*—but does not appear to have been followed up, either by Max Müller or his successors.

---

* Max Müller, 'Lectures on the Science of Language,' part 2, pp. 10, 89. See 'The Polynesian,' Brit. Mus. Lib., P.P. 9899b.

It has also been recognised[*] that it is the gestures of articulation, rather than the sounds, which are significant, and that many words, relating to actions of the organs of articulation, have been produced by the repetition of those actions during phonation. Thus the action of eating becomes the word mniʌm–mniʌn (mnyum–mnyum) when phonated.

The present observations extend this principle of word formation to cases where the tongue and lips copy the pantomimic movements of *other* organs, such as the human hands and arms.

It is difficult to suppose that the pantomimic gestures used by the original inventors of human speech could have survived unchanged up to the present or to historic times.

A consideration of the invented words used by children, indicates, however, that the tendency to make pantomimic gestures of articulation is, to some extent, innate. Thus, in Professor Otto Jespersen's "Language," at p. 152, he instances (*inter alia*) as children's invented words :—z–ɑ–ɑ, gön, and adi—which, if whispered on an in-breath, are found to be drinking, swallowing and tasting gestures, respectively. Similarly, fu'we (soap) may be due to the gesture of blowing away soapsuds, while dèdets (horse) seems due to a trotting or galloping gesture.

The present writer recently obtained particulars of the case of a boy, Grenville Gilbert, of Ware, Mass., U.S.A., who, between the ages of 3½ and 4 years, used an invented language of his own. In that language, the word for little was "ee–ee" (i–i), the word for big was "oh–oh" and the word for dog was "ba" (a barking gesture). It can hardly be a coincidence that this child's words for "big" and "little" were the same as those used by the Polynesians and ancient Japanese. The words referred to by Dr. Whymant may, therefore, have been re-invented, and need not necessarily be considered as unchanged descendants of the original words used by the first inventors of human speech.

[*Note added April,* 1928.—Since the paper was communicated, the following confirmations have been received from Dr. Whymant of the fabricated words appearing on page 158:

---

[*] Wilhelm Wundt, 'Völkerpsychologie,' vol. 1, pp. 116, 126, 136, 332 (Leipzig, 1904).

Action.	Nearest Phonetic Resultant.	Actual archaic Japanese, Polynesian and Oceanic.
Sew	PUI-SRI	TUI.
Roll or coil back	KƆIRA	KOERU. KIIRA.
Roll forward	TRIƆE TRIƆE	TRIE.
Blow through a tube	KY (as in French cul)	KI (-RI). KURU.
Dance up and down, to and fro	LI-LƆ	LI-LO LI-LŬ Indo-Chinese.
Plough	KƆGI	GAKE.
Strip grain off a stalk	THP or DHB	TUP or TUB (Indo-Chinese).
Bury in the ground	TVTH or TVS	TORO TANU

In addition Dr. Whymant has also given an actual word for shoot with a bow and arrow, to compare with the phonetic resultant on page 158, namely, THRKI; viz., KOKI or IKOKI. In this case the K of KOKI may represent a strongly reflexed R which ultimately becomes a species of K.]

The Relation of Articulation to Phonation.

It has already been pointed out that the speech sounds which are due to characteristic gestures of articulation are largely independent of the pitch of the laryngeal reed-note, by which the gesticulating cavities are made to resonate—indeed, the whole of the English speech sounds can be produced without laryngeal accompaniment, namely, in the form of whispered speech.* Phonation—*i.e.*, the range of sounds due to the vibratory action of the vocal chords—appears in fact to be a separate language, expressive of man's emotional state, and directly comparable with the emotional cries which are widely used throughout the animal world.

If the appearance of man's vocal chords could be observed during active change of his emotional state, it would probably be found that their "expression" changed quite as much as that of his lips, and that the various kinds of emotional cries were audible results of these changes of expression of the vocal chords and their surrounding parts, when energised by a flow of air from the lungs.

In voiced speech, the gesture language of articulation—which symbolises the ideas intended to be conveyed from the speaker to the listener—is combined more or less intermittently, with the expressional language of phonation, which indicates the speaker's emotional state.

---

* S.P.E. Tract No. 22, English Vowel Sounds, Oxford University Press.

In the so-called tone-languages, such as Chinese, phonation has been annexed to articulation. Even in English (as is well known) phonation has in some instances a verbal or grammatical significance—as when a speaker converts a sentence which is, in grammatical form, an assertion, into a question—by the expedient of raising the pitch of his voice at the end of the sentence, instead of lowering it, as is normally done. Except for this expedient, phonation, in English, is still essentially separate from articulation.

It has been pointed out by the present writer* that all language is still in a very primitive condition, and that the advantages of phonation, in giving increased range and musical value to speech, are not yet made full use of—as witness the presence in English speech of the unvoiced sounds, s, f, θ, and ʃ (sh). The consequent discontinuity of the emotional language may, it is suggested, be cited as an additional imperfection in our present languages.

The Action of the False Vocal Chords.

In a previous communication,† experiments were described which indicated that the essential difference between the so-called "voiced" and "unvoiced" consonants—*i.e.*, b, d, g, v, ð, z and ʒ, on the one hand, and p, t, k, f, θ, s, and ʃ, on the other—was due to a change in the form of the pharnyx, such as might be produced by the action of the false vocal chords. Confirmation of this theory has been obtained from Prof. G. Oscar Russell, of Ohio University, U.S.A. Prof. Russell has developed a pharyngeal periscope (a description of which is in preparation by him), by which the action of the pharynx and of the true and false vocal chords may be observed during articulation. Through this instrument it was possible to see his false vocal chords rising up and approaching one another when the consonant B was articulated (as compared with its attitude for the consonant P), so as to form a constriction comparable to that which was found to produce the "voiced" condition in the model. Acknowledgment should be made here of Dr. Russell's courtesy in permitting reference to his unpublished work.

It would appear that the false vocal chords must be definitely classed with the organs of articulation.

---

* S.P.E. Tract, No. 22.

† "The Nature and Artificial Production of the So-called Voiced and Unvoiced Consonants," 'Roy. Soc. Proc.,' A, vol. 114, pp. 98–102.

Electron Emission in Intense Electric Fields.

By R. H. Fowler, F.R.S., and Dr. L. Nordheim.

(Received March 31, 1928.)

§ 1. *Introduction.*—The main features of the phenomenon of the extraction of electrons from cold metals by intense electric fields are well known, and an approximate theory of the effect was first developed by Schottky.* More recently the experimental data have been much improved, notably by Millikan and Eyring,† and Millikan and Lauritsen‡. The theory has been considered afresh by O. W. Richardson§ and by Houston‖ working with Sommerfeld. It seems to us, however, that there is still room for improvement in the theoretical exposition and its correlation with the experiments. Neither O. W. Richardson nor Houston really treat the theory in the simple straightforward way which is now possible in the new mechanics, using the revived electron theory of metals which we owe to Sommerfeld. Again, while Millikan and Lauritsen seem to have established quite definitely the laws of dependence of the emission on the field strength F, they speak of the implications of their result in a way which is hard to justify and might in certain circumstances prove to be definitely misleading.

Millikan and Lauritsen show that a plot of log I, where I is the current, against 1/F yields a good straight line whenever the experimental conditions are sufficiently stable. At ordinary temperatures these currents are completely independent of the temperature. The formula for these currents is

$$\mathrm{I} = \mathrm{C}e^{-a/\mathrm{F}}, \tag{1}$$

which is, of course, indistinguishable from

$$\mathrm{I} = \mathrm{CF}^2 e^{-a/\mathrm{F}}. \tag{2}$$

Millikan and his associates have also shown that as the higher temperatures, at which ordinary thermionic emission begins, are approached, the strong field emission does become sensitive to temperature and finally blends into the thermionic.

---

* Schottky, 'Z. f. Physik,' vol. 14, p. 80 (1923).

† Millikan and Eyring, 'Phys. Rev.,' vol. 27, p. 51 (1926).

‡ Millikan and Lauritsen, 'Proc. Nat. Ac. Sci.,' vol. 14, p. 45 (1928).

§ O. W. Richardson, 'Roy. Soc. Proc.,' A, vol. 117, p. 719 (1928).

‖ Houston, 'Z. f. Physik,' vol. 47, p. 33 (1928).

On the strength of these facts they suggest that perhaps there may exist a general formula for the current

$$I = A\,(T + cF)^2\, e^{-b/(T+cF)}, \tag{3}$$

valid over wide ranges of temperature (T) and field strength. This formula is, of course, correct for large T and small F and also for large F and small T. At intermediate strengths it does not appear to have been tested quantitatively and therefore yet awaits experimental and theoretical investigation. On one other deduction, however, they lay great emphasis, and this we find liable to mislead. They assert that a distinction should be drawn between the electrons which can function as thermions and the ordinary conduction electrons which yield the emission at great field strengths and are absolutely independent of the temperature.

In this paper, therefore, we extend the results of Nordheim[*] to include the effect of an external field using the same methods and the same underlying picture of the metal (Sommerfeld's). We establish the formula (2) independent of T at low temperatures in agreement with experiment. We fail to find any theoretical justification for (3), though, of course, some justification may exist. Combining our results with those of Nordheim we show that Sommerfeld's picture of a metal yields the formula both for strong fields and for thermionic emission. A single set of free or conduction electrons distributed according to the Fermi-Dirac statistics suffices for both purposes. It is for this reason that we take exception to the statements of Millikan and Lauritsen recorded above. We gladly admit that everything they say can be reconciled with our theory, but only by a forced interpretation of their statements which we consider lay the emphasis wrongly.

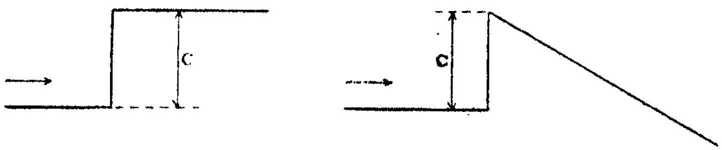
Our calculations are closely allied to some recent work by Oppenheimer[†] on the hydrogen atom in an external electric field. Oppenheimer notices that his work has a bearing on the emission of electrons in strong fields, but he does not pursue the matter further. For our purposes the calculations can be shorn of irrelevancies and made so much simpler that it is worth while attacking the problem *de novo*.

§ 2. *The Reflection of Electrons at a Potential Jump when an Electric Field acts on one Side.*—In the paper quoted Nordheim has calculated exactly the reflection or emission coefficients for electrons of given energy W incident from the left on a surface at which their potential energy is suddenly increased by an amount C, fig. 1 (i). The problem is one-dimensional, and W is the kinetic energy of the electron's motion normal to the surface of the metal. The rest

---

[*] Nordheim, 'Z. f. Physik,' vol. 46, p. 833 (1928).

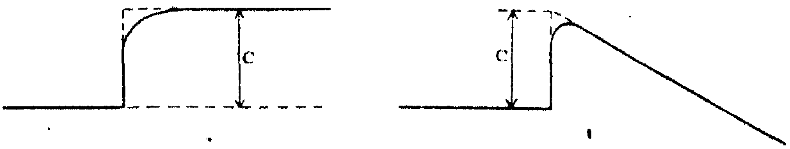
[†] Oppenheimer, 'Phys. Rev.,' vol. 13, p. 66 (1928).



(i) (ii)

Fig. 1.

of its kinetic energy is irrelevant to the problem. He has shown further that the results are only slightly modified by modifying the form of the step, *e.g.*, by slightly rounding off the corners. We shall here calculate in a similar way the emission coefficients for electrons of given energy when a uniform external field acts, so that the potential energy of the electrons is as shown in fig. 1 (ii). The corner at the top will, of course, really be rounded off in both cases by the image effect. This will seriously alter the emission coefficient for non-zero external field for electrons of incident energy nearly equal to W. We shall see, however, that this modification is unimportant in calculating the strong field emission at ordinary temperatures. The potential energies rounded off by the image effect will be somewhat as shown in fig. 2.



(i) (ii)

Fig. 2.

In order to study the emission through the potential energy step of fig. 1 we have only to solve the wave equations

$$\frac{d^2\psi}{dx^2} + \kappa^2 (W - C + Fx)\,\psi = 0 \qquad (x > 0), \tag{4}$$

$$\frac{d^2\psi}{dx^2} + \kappa^2 W\psi \qquad\qquad = 0 \qquad (x < 0), \tag{5}$$

subject to the conditions that ψ and $d\psi/dx$ are continuous at $x = 0$ and that for $x > 0$ ψ represents a stream of electrons progressing to the right only. The constant κ is defined by

$$\kappa^2 = 8\pi^2 m/h^2. \tag{6}$$

At all ordinary and low temperatures practically all the electrons are so distributed according to Sommerfeld's formula (see equation (19)) that $W < C$, so that we have to discuss the case in which we pass through a zero of the coefficient of ψ.

Write

$$\left(-\frac{C-W}{F}+x\right)(\kappa^2 F)^{\frac{1}{3}} = y.$$

then equation (4) becomes

$$\frac{d^2\psi}{dy^2}+y\psi = 0,$$

of which the solutions are expressible in terms of Bessel's functions* of order $\frac{1}{3}$,

$$\psi = \sqrt{y}\,J_{\pm\frac{1}{3}}\left(\tfrac{2}{3}y^{\frac{3}{2}}\right). \tag{7}$$

We require that solution which for large x (*i.e.*, y) represents a wave travelling to the right. Therefore

$$\psi = \sqrt{y}\,H_{\frac{1}{3}}^{(2)}\left(\tfrac{2}{3}y^{\frac{3}{2}}\right), \tag{8}$$

where $H^{(2)}$ denotes the second function of Hankel. For large y

$$\psi \sim \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{1}{y^{\frac{1}{4}}}\frac{1}{(2/3)^{\frac{1}{2}}}\,e^{-i\left[\frac{2}{3}y^{\frac{3}{2}}-\frac{5\pi}{12}\right]}. \tag{9}$$

Observe that as $y \longrightarrow \infty$

$$|\,\psi\,|^2 \sim \frac{A}{(W-C+Fx)^{\frac{1}{2}}} = \frac{A'}{v},$$

where A, A' are constants and v is the velocity of the electrons. Hence the density of the electron stream behaves as it should.

For $x < 0$ we take

$$\psi = \frac{1}{W^{\frac{1}{4}}}\left[ae^{-i\kappa x\sqrt{W}}+a'e^{-i\kappa x\sqrt{W}}\right]. \tag{10}$$

We have now to study what happens to $H_{\frac{1}{3}}^{(2)}$ as we pass back to $x = 0$ from large x through the zero of y. There is no real ambiguity, for ψ is represented by a power series of integral powers of y. In order, however, to use the normal asymptotic expansion for $H_{\frac{1}{3}}^{(2)}\left(\frac{2}{3}y^{\frac{3}{2}}\right)$, which is valid in the range

$$-2\pi < \arg\left(y^{\frac{3}{2}}\right) < \pi,$$

we must take for negative y

$$y = -y' = y'e^{-\pi i},$$

---

* For all formulæ used here involving Bessel's functions, see Watson, "Bessel's Functions." They are easily found so that it is unnecessary to give detailed references.

and this convention we shall use throughout. At $x = +0$ we then find

$$\psi(0) = \left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac12} e^{-\frac12\pi i}(\kappa^2\mathrm{F})^{\frac16}\,\mathrm{H}^{(2)}_{\frac13}\left(e^{-\frac32\pi i}\tfrac23\kappa\sqrt{\mathrm{F}}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac32}\right), \tag{11}$$

$$\left(\frac{d\psi}{dx}\right)_0 = \tfrac12\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{-\frac12} e^{\frac12\pi i}(\kappa^2\mathrm{F})^{\frac16}\,\mathrm{H}^{(2)}_{\frac13}\left(e^{-\frac32\pi i}\tfrac23\kappa\sqrt{\mathrm{F}}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac32}\right)$$
$$-\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac12}(\kappa^2\mathrm{F})^{\frac16}\,\kappa\,(\mathrm{C}-\mathrm{W})^{\frac12}\,\mathrm{H}^{(2)\prime}_{\frac13}\left(e^{-\frac32\pi i}\tfrac23\kappa\sqrt{\mathrm{F}}\left(\frac{(\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac32}\right). \tag{12}$$

These values are to be equated to the values of ψ and $\partial\psi/\partial x$ for $x = -0$ derived from (10).

Since the only quantity of physical interest is $|a|^2/|a'|^2$, which determines the emission coefficient, we can simplify the resulting equations by omitting any common factor real or complex which is thereby absorbed in both a and a'. We shall write also

$$\tfrac23\kappa\sqrt{\mathrm{F}}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac32} = \mathrm{Q},$$

so that Q is real and in practice large. Further,

$$\mathrm{H}^{(2)\prime}_{\frac13}(e^{-\frac32\pi i}\mathrm{Q}) = e^{\frac32\pi i}\frac{d\mathrm{H}^{(2)}_{\frac13}}{d\mathrm{Q}}(e^{-\frac32\pi i}\mathrm{Q}).$$

The equations of continuity of ψ and $d\psi/dx$ can therefore be reduced to

$$a + a' = \mathrm{W}^{\frac14}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac12}\mathrm{H}^{(2)}_{\frac13}(e^{-\frac32\pi i}\mathrm{Q}), \tag{13}$$

$$-a + a'$$
$$= \frac{i}{\kappa\mathrm{W}^{\frac14}}\left[\tfrac12\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{-\frac12}\mathrm{H}^{(2)}_{\frac13}(e^{-\frac32\pi i}\mathrm{Q}) + \frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\,\kappa\sqrt{\mathrm{F}}\,\frac{d\mathrm{H}^{(2)}_{\frac13}}{d\mathrm{Q}}(e^{-\frac32\pi i}\mathrm{Q})\right]. \tag{14}$$

By the definition of the functions of Hankel we can express $\mathrm{H}^{(2)}_{\frac13}(e^{-\frac32\pi i}\mathrm{Q})$ in terms of the real functions $\mathrm{I}_{\pm\frac13}(\mathrm{Q})$ by the equation

$$\mathrm{H}^{(2)}_{\frac13}(e^{-\frac32\pi i}\mathrm{Q}) = \frac{-1}{\sin\frac13\pi}\{\mathrm{I}_{-\frac13}(\mathrm{Q}) + e^{\frac13\pi i}\,\mathrm{I}_{\frac13}(\mathrm{Q})\}. \tag{15}$$

Let us now write

$$\alpha + i\beta = \frac{\mathrm{I}'_{-\frac13} + e^{\frac13\pi i}\,\mathrm{I}'_{\frac13}}{\mathrm{I}_{-\frac13} + e^{\frac13\pi i}\,\mathrm{I}_{\frac13}}, \tag{16}$$

where α and β are real. Let us also write D (W) for the fraction of W electrons penetrating the boundary peak and emerging under the influence of the external field F. Then by solving (13) and (14) and using (15) and (16) we find

$$\mathrm{D}\,(\mathrm{W}) = \frac{|a|^2 - |a'|^2}{|a|^2}$$

$$= \frac{4\beta\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac{3}{2}}\sqrt{\mathrm{F}}}{\left\{\mathrm{W}^{\frac{1}{4}}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac{1}{2}} + \frac{\mathrm{C}-\mathrm{W}}{\sqrt{\mathrm{F}}}\,\frac{\beta}{\mathrm{W}^{\frac{1}{4}}}\right\}^2 + \frac{1}{\kappa^2\mathrm{W}^{\frac{1}{2}}}\left\{\frac{1}{2}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{-\frac{1}{2}} + \frac{\mathrm{C}-\mathrm{W}}{\sqrt{\mathrm{F}}}\,\kappa\alpha\right\}^2}.$$

It remains to calculate α and β. It is easily verified that

$$\beta = \frac{\sqrt{3}}{2}\,\frac{(\mathrm{I}_{\frac{1}{3}}'\,\mathrm{I}_{-\frac{1}{3}} - \mathrm{I}_{-\frac{1}{3}}'\,\mathrm{I}_{\frac{1}{3}})}{(\mathrm{I}_{-\frac{1}{3}} + \frac{1}{2}\mathrm{I}_{\frac{1}{3}})^2 + \frac{3}{4}\mathrm{I}_{\frac{1}{3}}^{2}}.$$

The numerator is the Wronskian of Bessel's equation of purely imaginary argument and we have exactly

$$\mathrm{I}_{\frac{1}{3}}\,\mathrm{I}_{-\frac{1}{3}}' - \mathrm{I}_{-\frac{1}{3}}\,\mathrm{I}_{\frac{1}{3}}' = -\frac{2\sin\frac{1}{3}\pi}{\pi\mathrm{Q}}.$$

For the denominator we can use the asymptotic expansion, Q being large. The denominator is

$$|\sin\tfrac{1}{3}\pi\,\mathrm{H}_{\frac{1}{3}}^{(2)}(e^{-\frac{1}{2}\pi i}\,\mathrm{Q})|^2 \sim \tfrac{3}{4}\,\frac{2}{\pi\mathrm{Q}}\,e^{2\mathrm{Q}},$$

so that

$$\beta \sim e^{-2\mathrm{Q}}.$$

In evaluating α we can use asymptotic values throughout and find $\alpha = 1$. Thus with sufficient accuracy

$$\mathrm{D}\,(\mathrm{W}) = \frac{4\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{\frac{3}{2}}\sqrt{\mathrm{F}}e^{-2\mathrm{Q}}}{\mathrm{W}^{\frac{1}{2}}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right) + \frac{1}{\kappa^2\mathrm{W}^{\frac{1}{2}}}\left\{\frac{1}{2}\left(\frac{\mathrm{C}-\mathrm{W}}{\mathrm{F}}\right)^{-\frac{1}{2}} + \frac{\mathrm{C}-\mathrm{W}}{\sqrt{\mathrm{F}}}\,\kappa\right\}^2}. \tag{17}$$

By considering the relative numerical order of the various terms in the denominator it is at once found that those independent of κ are dominant. Therefore we find with sufficient accuracy on inserting the value of Q

$$\mathrm{D}\,(\mathrm{W}) = \frac{4\,\{\mathrm{W}\,(\mathrm{C}-\mathrm{W})\}^{\frac{1}{2}}}{\mathrm{C}}\,e^{-4\kappa(\mathrm{C}-\mathrm{W})^{\frac{3}{2}}/3\mathrm{F}}. \tag{18}$$

§ 3. *The Complete Electron Emission from a Cold Metal.*—The number of electrons N (W) incident on a surface of unit area per unit time with a kinetic energy W normal to the surface has been evaluated by Nordheim[*] according to Sommerfeld's theory. He finds

$$\mathrm{N}\,(\mathrm{W}) = \frac{4\pi mk\mathrm{T}}{h^3}\,\mathrm{L}\left(\frac{\mathrm{W}-\mu}{k\mathrm{T}}\right), \tag{19}$$

where

$$\mathrm{L}\,(\beta) = \int_0^\infty \frac{dy}{e^{\beta+y}+1}, \tag{20}$$

and μ is the usual parameter of the electron distribution in the Fermi-Dirac statistics equivalent to the thermodynamic partial potential of an electron. Hence the current I is given quite generally by

$$\mathrm{I} = \frac{4\pi m\varepsilon k\mathrm{T}}{h^3}\int_0^\infty \mathrm{D}\,(\mathrm{W})\,\mathrm{L}\left(\frac{\mathrm{W}-\mu}{k\mathrm{T}}\right)d\mathrm{W}, \tag{21}$$

where ε is the electronic charge. For all low and ordinary temperatures the integrand of $\mathrm{L}\,\{(\mathrm{W}-\mu)/k\mathrm{T}\}$ is only sensible when $(\mathrm{W}-\mu)/k\mathrm{T}+y<0$ and is then unity. Thus at such temperatures a sufficient approximation to $k\mathrm{T}\,\mathrm{L}\,\{(\mathrm{W}-\mu)/k\mathrm{T}\}$ is $\mu-\mathrm{W}$ when $\mathrm{W}<\mu$ and otherwise zero. The reason for the sufficiency of this approximation is that in actual cases W and μ are in general very large compared with kT. Since μ is considerably less than C, we may then use (18) for D (W) and find

$$\mathrm{I} = \frac{16\pi m\varepsilon}{\mathrm{C}h^3}\int_0^\mu \mathrm{W}^{\frac{1}{2}}\,(\mathrm{C}-\mathrm{W})^{\frac{1}{2}}\,(\mu-\mathrm{W})\,e^{-4\kappa(\mathrm{C}-\mathrm{W})^{\frac{3}{2}}/3\mathrm{F}}\,d\mathrm{W}.$$

Since the exponent in the integrand is still very large for the largest values of W, it is easy to evaluate this integral to a sufficient approximation. We find, using (6) and putting $\mathrm{C}-\mu=\chi$, that

$$\mathrm{I} = \frac{\varepsilon}{2\pi h}\,\frac{\mu^{\frac{1}{2}}}{(\chi+\mu)\,\chi^{\frac{1}{2}}}\,\mathrm{F}^2 e^{-4\kappa\chi^{\frac{3}{2}}/3\mathrm{F}}. \tag{21}$$

The χ of this equation is necessarily and exactly the thermionic work function.[†]

If, for convenience of discussion, we express I in amperes per square centimetre of emitting surface, μ and χ in volts and F in volts per centimetre, and insert numerical values for the other constants, we find

$$\mathrm{I} = 6\cdot 2\times 10^{-6}\,\frac{\mu^{\frac{1}{2}}}{(\chi+\mu)\chi^{\frac{1}{2}}}\,\mathrm{F}^2 e^{-2\cdot 1\times 10^{8}\chi^{\frac{3}{2}}/\mathrm{F}}. \tag{22}$$

We are now in a position to compare the theory with the experimental facts. We see at once that I is of the correct form by comparison with (1) or (2). We have not calculated explicitly the temperature effect on (22), but it is easy to see that it is very small at ordinary temperatures. Formula (22) refers strictly to the limit $T = 0$, but it is a valid approximation so long as, let us say, μ/kT is very large. Now μ is of the order of 5 volts and kT is $8 \cdot 6 \times 10^{-5}T$ in the same units. This is sufficient to guarantee the observed independence of T for all ordinary temperatures.

---

[*] Nordheim, *loc. cit.*, formula (11).

[†] Fowler, 'Roy. Soc. Proc.,' A, vol. 117, p. 549 (1928).

The form of the exponent $\chi^{\frac{3}{2}}/F$ is interesting, but there seems no possibility at present of any experimental test of the form.

We come now to the absolute value of the exponent. For the metals commonly experimented with we may take $\chi^{\frac{3}{2}} = 10$, and we can take the exponential factor nearly enough to be $10^{-10^9/F}$. This will make the emission begin to be sensible for fields of rather more than 10^7 volts/cm. The emission would be already very large for fields of the order 10^8 volts/cm. These values are higher than those commonly derived from experiments, which are commonly taken to indicate measurable emission for values of F about 10^6 volts/cm. The absolute determination of F is, however, very difficult, owing to the very large effect of minute surface irregularities or peaks near which we shall find larger values of F than those derived from the geometry of the apparatus. There is also the possibility of sensitive spots on the surface with a reduced value of χ. An investigation by Rother* seems to support this view, for he finds that after very careful preliminary heat treatment an F as large as 10^7 volts/cm. is required to extract a reasonable current. It seems fair to conclude that the phenomenon of electron emission in intense fields is yet another phenomenon which can be accounted for in a satisfactory quantitative manner (at least in broad outline) by Sommerfeld's theory.

§ 4. *Additional Calculations.*—It is not difficult to show generally, by comparison with Jeffreys'† asymptotic solution of a similar equation, that for values of W not too near C the exact form of the potential peak and the rounding off at the top will not seriously affect the emission coefficient D (W). Our direct calculation for this very simple case is therefore sufficiently typical. For the Q of an exact solution has to be replaced by

$$\kappa \int_{x_0}^{x_1} (V - W)^{\frac{1}{2}} \, dx,$$

where V is the potential energy of the electron at any point and x_0 and x_1 are the points at which $V - W$ vanishes. The integration range is shown in fig. 3. It is at once clear that, provided the shaded area is reasonably large, modifications in the contour near the peak are unimportant.

---

* Rother, 'Ann. der Physik,' vol. 81, p. 316 (1926).

† Jeffreys, 'Proc. Lon. Math. Soc.,' vol. 23, p. 428 (1924).

We pass on to an outline of the calculations necessary to establish the form of the combined thermionic and strong field emission. In the first place, if

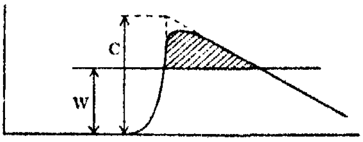


FIG. 3.

$W > C$, the calculation of D (W) can be carried out exactly as in Nordheim's paper, and we find

$$D(W) = \frac{4W^{\frac{1}{2}}(W-C)^{\frac{1}{2}}}{[W^{\frac{1}{2}} + (W-C)^{\frac{1}{2}}]^2 + [F/(4\kappa\{W-C\})]^2}. \tag{23}$$

From this formula it is easily verified that the term in F is quite unimportant in ordinary cases, and the emission can be calculated as if $F = 0$.

Near the peak itself, when W is nearly equal to C, neither formula (18) nor (23) can be strictly used, and D (W) will depend essentially on the exact form of the potential energy curve. All we can say at present is $D(W) = 0$ when $W = V_{\text{max.}}$, the maximum ordinate, at which in general the tangent will be horizontal, and that for other neighbouring values of W D (W) is of the order of $\frac{1}{2}$.

We can now break up the range of integration of W in the general formula (21) into four parts as in fig. 4. We can see at once that of the total emission

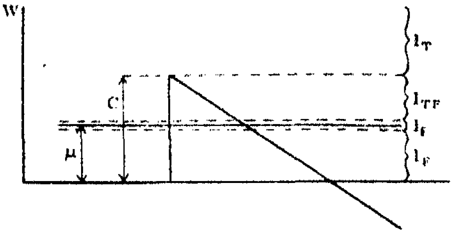


FIG. 4.

I_T and I_F represent fairly accurately the purely thermionic and purely strong field emission respectively. In I_ϵ (which, however, is unimportant), a more accurate value of L $\{(W-\mu)/kT\}$ must be used than in I_F, and both I_ϵ and I_{TF} depend essentially on both T and F. Calculations for the field of fig. 1 are not of sufficient importance to give in detail, but they give no combined result of the form (3) nor any grounds for expecting such a result for any natural modification of the field.

On Line-Integrals of the Diurnal Magnetic Variations.

By A. T. Price, M.Sc., and S. Chapman, F.R.S.

(Received March 10, 1928.)

Introduction.

1. According to electromagnetic theory, the line-integral $\int \mathbf{H} \cdot d\mathbf{s}$ of the magnetic force H taken round any closed curve is equal to $4\pi I$, where I is the electric current threading the curve, H and I being measured in c.g.s. units. Such line-integrals have been calculated by Gauss* and many later investigators† for various curves on the earth's surface, in order to determine whether any electric current flows upwards or downwards across the surface. Modern computations for large areas lead in general to values of $\int \mathbf{H} \cdot d\mathbf{s}$ differing from zero by amounts that correspond to current-densities of the order $3 \cdot 10^{-2}$ ampere/km.^{2}. The magnetic field of such currents would account for 2 or 3 per cent. of the earth's surface field.

These results are inconsistent with the direct measurements of the atmospheric electric potential gradient and the ionisation of the air, which indicate a vertical-current-density of the order $3 \cdot 10^{-6}$ amp./km.^{2}. If the magnetic estimates are reliable, the discrepancy indicates either that atmospheric electric currents exist which escape measurement, though they are 10,000 times as great as those which are measured, or that the relation $\int \mathbf{H} \cdot d\mathbf{s} = 4\pi I$, which is one of the foundations of electromagnetic theory, is not strictly correct. These alternatives are so remarkable that the magnetic evidence must be above suspicion if it is to gain credence. Dr. L. A. Bauer holds that the results got from independent sets of data, for different epochs, and the mutual accordance of the results from neighbouring areas, justify the acceptance of the non-zero line-integrals, and that to explain them away it is necessary to assume quite unlikely systematic errors in the magnetic data. Other investigators show less conviction: for example, Sir Frank Dyson and H. Furner conclude that "though there is some evidence for Prof. Bauer's results, the existence of vertical electric currents is not indicated with any great certainty." But though the magnetic evidence may not be conclusive, it cannot be lightly dismissed, and in view of the importance of the question Sir Arthur Schuster* has recently urged the desirability of a detailed magnetic survey of a small area as the best means of obtaining a definite conclusion.

---

* Gauss, "Allgemeine Theorie der Erdmagnetismus," p. 13.

† Schmidt, 'Münchener Ak. Abh.,' Kl. 2, vol. 19 (1895); 'Archiv. d. D. Seewarte,' vol. 21 (1898); 'Z. f. Geophysik,' vol. 1, p. 1 (1925). Rücker, 'Terr. Mag.,' vol. 2, p. 84 (1896); 'Nature,' vol. 57, pp. 160, 180, where further references are given. Bauer, 'Terr. Mag.,' vol. 2, p. 11 (1897); vol. 9, p. 116 (1904); vol. 25, p. 345 (1920); March-June, 1923; Bauer, L. A., and Peters, W. J., 'Terr. Mag.,' Sept., 1923. Dyson and Furner, 'M.N.R.A.S.,' Geoph. Supp. 1, p. 76 (1923).

If there is no electric earth-air current, the magnetic field can be expressed as the gradient of a scalar potential. The attempts to determine such a potential function by means of spherical harmonic analysis have not been completely successful. Dr. Adolf Schmidt (*loc. cit.*) modified the analysis of Gauss to allow for the possibility of a part of the field being due to earth-air currents, and his computations indicated the presence of such a part, corresponding to an average vertical-current-density of about $15 \cdot 10^{-2}$ amp./km.^{2}. But the estimates derived later by means of line-integrals, and the discrepancy with atmospheric-electric observations, led him to doubt the reliability of the magnetic evidence.

2. Similar questions arise in connection with the diurnal variation of the earth's magnetic field. Schuster† and many later writers have attempted to express this varying part of the field by means of a scalar potential, but their success has been only moderate and incomplete. This may well be due to the irregularity in the distribution of the field, and the limited extent of the available data, which restricts the number of harmonic terms that can usefully be included in the analysis; moreover, the data themselves are subject to certain errors. But Miss van Vleuten,‡ who made such an analysis, concluded that a considerable part of the diurnal-variation field does not possess a potential, and is therefore due to vertical currents. In a critical review§ of her work, Dr. Adolf Schmidt showed that the apparent failure of her expressions for the potential arose almost wholly in the southern hemisphere, where no observational data had been used, and that in the region for which data had been used in her work the conditions for the validity of a potential expression were satisfied with reasonable accuracy; he concluded that it was safe to assert the irrotational character of the horizontal field of the daily magnetic variation to a degree of accuracy far above that of the magnetic data yet available.

* 'Roy. Soc. Proc.,' A, vol. 111, p. 68 (1926).

† 'Phil. Trans.,' A, vol. 180, p. 467 (1889).

‡ 'K. Nederl. Met. Inst.,' Bulletin 102 (1917).

§ 'Phys. Z.,' vol. 19, p. 349 (1918).

Dr. Schmidt's arguments, which appear to us convincing, are not universally accepted. Dr. Bauer[*] adopts Miss van Vleuten's conclusion that there are diurnally varying earth-air currents, the magnitude of the range of variation being about $3 \cdot 10^{-3}$ amp./km.^{2}, or about one-tenth as great as the *mean* current-densities estimated on magnetic grounds (§ 1). Using Miss van Vleuten's formulæ, he computed the daily summer variation in i, the current-density, for a station in latitude 39° N., assuming that i is the same, at the same local times, at all points on any circle of latitude. His curve (reproduced here in fig. 2, curve a) proves to be similar in form to that of the diurnal variation of i as estimated by atmospheric-electric measurement, at Potsdam (52° N.) and Davos (47° N.), but its amplitude is 1000 times as great. The observed variations of i are much too small to affect the daily magnetic variations appreciably.

In view of the difference of opinion which exists on this question, we decided to examine it by the most direct method, namely, by calculation of line-integrals of the daily magnetic variation; it is somewhat surprising that the present is, so far as we are aware, the first occasion on which this has been done.

Line-Integrals of the Daily Magnetic Variation.

3. Let $\overline{\mathrm{H}}$ denote the mean horizontal magnetic force-vector over a period of a day (or many days), and $\Delta\mathrm{H}$ the difference from this mean at any particular hour of the day. Considering any closed curve,

$$\int \mathbf{H} \,.\, \mathbf{ds} = \int \overline{\mathbf{H}} \,.\, \mathbf{ds} + \int \Delta\mathbf{H} \,.\, \mathbf{ds}$$

$$= 4\pi\overline{\mathrm{I}} + 4\pi\Delta\mathrm{I} = 4\pi\mathrm{A}\,(\overline{i} + \Delta i),$$

where $\overline{i}$ is the mean current density over the enclosed area A, and Δi is the difference from this mean at the given hour.

In this paper we shall not be concerned with $\overline{i}$, but only with $\int \Delta\mathbf{H} \,.\, \mathbf{ds}$ and Δi. Several circumstances favour the accuracy of the calculation of Δi as compared with that of $\overline{i}$: (i) $\Delta\mathrm{H}$ is continuously and simultaneously registered at permanent magnetic observatories much more exactly than **H** can be measured in magnetic surveys; (ii) no reduction to a common epoch is required; (iii) average values over a considerable number of quiet days can be used; and (iv) $\Delta\mathrm{H}$ is perhaps less subject to local modifications than is **H** itself. Moreover, Δi can be calculated for each of 24 successive hours, and if the values represent real currents the variation from hour to hour should be systematic.

[*] 'Terr. Mag.,' vol. 25, p. 160 (1920).

But to set against these favourable circumstances there is the fact that magnetic surveys give values of H for a much closer network of points than that formed by magnetic observatories. The great distance between neighbouring observatories proves, in fact, to be the chief difficulty in the present investigation, and it seems possible to meet it only by the same method as that proposed by Schuster for the main field, namely, by special observations at stations closer together; this, however, would be an expensive and elaborate undertaking.

4. After a preliminary examination of the available data we decided to deal with a curve in Europe, where magnetic observatories are least thinly spread. It seemed desirable to choose a curve having its greatest extension nearly transverse to the magnetic meridian, for with this choice $\int \Delta\mathbf{H} \, . \, d\mathbf{s}$ will depend more on observations of ΔD (where D denotes the magnetic declination) than on observations of the horizontal intensity H, and ΔD, being registered by a freely suspended magnet, with a scale value determined by purely geometrical measurements, is easier to measure accurately than ΔH, which is registered by an instrument whose scale value depends on the torsion of a fibre, and is affected by variations of temperature. It is true that the values of ΔD have to be multiplied by H to reduce them to force-units, but the absolute value of H at magnetic observatories is known with an accuracy amply sufficient for this purpose.

On this account, and for other reasons, the observatories of Greenwich, Eskdalemuir, Rude Skov, Seddin and de Bilt were chosen to mark the limits of the curve. Afterwards, by the courtesy of the Rev. Father Rowland, S.J., unpublished data for Stonyhurst were obtained and incorporated in part of the calculations. An increase in the area bounded by the curve would have increased the accuracy of the determinations of Δi, if the curve were of equally suitable form, with reliable observatories spaced at similar intervals along it; but these conditions are difficult to fulfil outside Europe. The boundary might have been enlarged to include further European stations, but the adopted choice was considered sufficient to indicate the present possibilities of the method.

5. The following table gives various relevant particulars for the six observatories: fig. 1 illustrates their positions, approximately to scale, and also the form of the two curves for which line-integrals have been computed.

Table I.

—	Greenwich. (G.)		Eskdale-muir. (E.)		Rude Skov. (R.)		Seddin. (S.)		de Bilt. (dB.)		Stony-hurst. (St.)	
Latitude	51° 28′		55° 19′		55° 51′		52° 17′		52° 6′		53° 51′	
Longitude	0		3° 12′ W.		12° 27′ E.		13° 1′ E.		5° 11′ E.		2° 28′ W.	
Elements recorded	N.	D.	N.	W.	H.	D.	N.	W.	H.	D.	H.	D.
Scale : γ/mm.	3	3	5	5	10	5	2	2	3	5	4	5
Temperature coefficient : $\gamma/^\circ$ C.	2	—	−9	−2	?	—	0·0	0·4	6	—	?	—

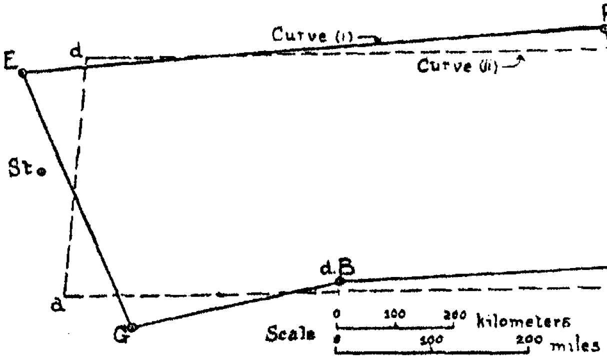


FIG. 1.

The extent of the curves is over 200 miles from N. to S., and about 500 miles from E. to W. Curve (1) consists of the five arcs of great circles joining the five observatories; its area is found from the formula $R^2\Sigma E$, where R denotes the earth's radius and E refers to the stations at the ends of each arc, and for a pair of stations (1, 2) is given by

$$\tan \tfrac{1}{2}E = \frac{\tan \tfrac{1}{2}\psi_1 \tan \tfrac{1}{2}\psi_2 \sin (\lambda_2 - \lambda_1)}{1 + \tan \tfrac{1}{2}\psi_1 \tan \tfrac{1}{2}\psi_2 \cos (\lambda_2 - \lambda_1)},$$

λ, ψ being the longitudes and co-latitudes. It appears that $A = 0{\cdot}00975\ R^2$, or 396,000 km.^{2}.

6. The data used were the average hourly values of ΔH and ΔD, or ΔN and ΔW (north and west components of force), for the 20 international quiet days of the summer (May to August, inclusive) of the year 1924, which was one of low sunspot number. The values represent differences from the mean of the 24 hourly measures from 1 h. to 24 h., Greenwich mean time; but 25 hourly differences were considered, for the hours 0 to 24, the first of these being included so as to indicate if possible any non-cyclic change in the supposed vertical

currents. No non-cyclic correction was applied, and where such had already been applied in the published data it was removed.

Usually the published data consist of hourly deviations from the mean, and are given to 1 γ; in obtaining the 20-day averages the deviations were computed to the nearest $0 \cdot 1\ \gamma$ in such a way that the sum of the last 24 did not exceed $1 \cdot 2\ \gamma$; this was convenient as an aid in checking the later work, though the average values may not be reliable to $0 \cdot 1\ \gamma$.

The published values for Seddin and Greenwich are mean values for periods of an hour centred at the half hours of Greenwich time. Graphs of the average diurnal variations ΔN and ΔW for these observatories were drawn to a scale of 1 cm. = 1 γ, and values at the exact hours were read off to $0 \cdot 1\ \gamma$. An examination of the Seddin data, which also include instantaneous values at hours of local time, indicated that (for quiet-day averages) this procedure introduced no appreciable error. Likewise it was considered unnecessary to make any allowance for the difference between the instantaneous hourly values for Eskdalemuir and the published hourly means centred at the exact hours. The data for Rude Skov and de Bilt are instantaneous hourly values.

7. In calculating $\int \Delta \mathbf{H} \, . \, d\mathbf{s}$ for curve (1), consisting of the great-circle arcs joining the five observatories, ΔN and ΔW were treated as constant along each arc, and as being equal to the mean of the values, ΔN or ΔW, for the two stations at the ends of the arc.

Thus $\int \Delta \mathbf{H} \, . \, d\mathbf{s}$, which is equal to $\int \Delta N dn + \int \Delta W dw$, where dn and dw denote the components of $d\mathbf{s}$ along the N. and W. directions, was taken to be given by $\Sigma\ (n \Delta N + w \Delta W)$, where n denotes the difference of latitude between the ends of the arc, reckoned in linear measure, and w is estimated, without appreciable error, as equal to $\frac{1}{2}R\ (\lambda_1 - \lambda_2)\ (\sin \psi_1 + \sin \psi_2)$. The possible error involved in this procedure is examined in §§ 9–11.

The values of n/R and w/R for the five arcs, starting from Greenwich and proceeding in the anti-clockwise direction (as viewed from above), are as follows :—

	G.—d.B.	d.B.—S.	S.—R.S.	R.S.—E.	E.—G.
n/R ..	0·0110	0·0031	0·0623	−0·0092	−0·0672
w/R .	−0·0561	−0·0839	0·0058	0·1544	−0·0334

The values of the 25 line-integrals, one for each hour, expressed in terms of the mean inferred upward current-density (obtained on dividing $\int \Delta \mathbf{H} \, . \, d\mathbf{s}$ by $4\pi A$) are given in Table II.

Table II.—Mean inferred upward current density Δi at each hour, from curve (1). Unit: 10^{-4} ampere/km.^{2}.

Hour.	Δi.	Hour.	Δi.	Hour.	Δi.	Hour.	Δi.
0	27	7	− 2	13	−25	19	16
1	9	8	− 3	14	6	20	6
2	− 2	9	−11	15	− 8	21	21
3	−11	10	−11	16	− 7	22	25
4	− 6	11	−41	17	− 4	23	22
5	− 2	12	−25	18	1	24	23
6	− 2	—	—	—	—	—	—

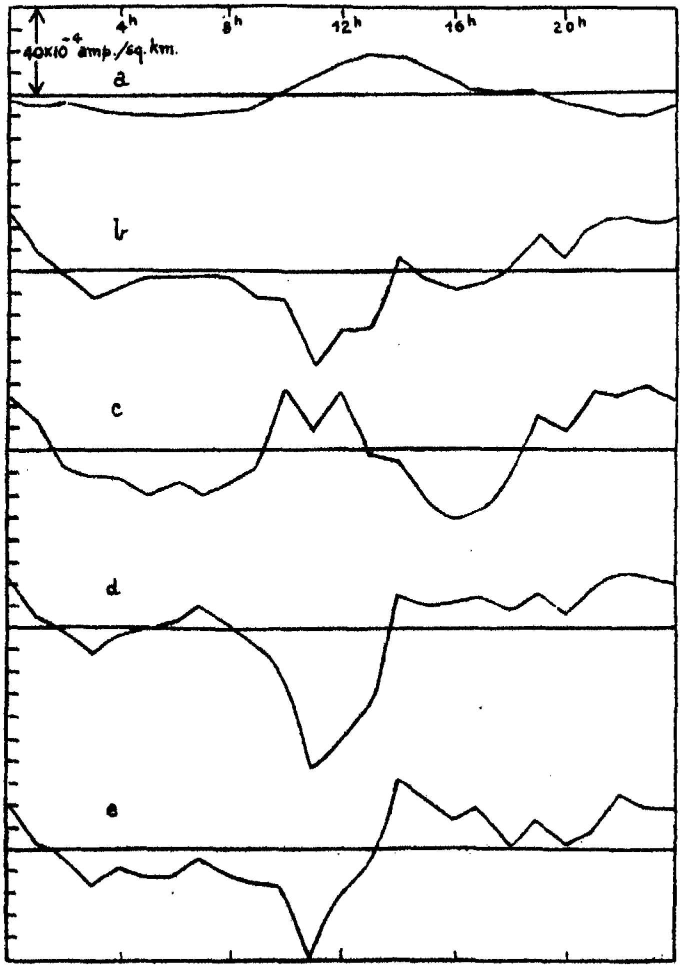


Fig. 2.

The maximum variation from the mean current-density is thus 41 . 10^{-4} amp./km.^{2}, at 11 h.; the range about the mean is 68 . 10^{-4} amp./km.^{2}, and the mean deviation from the mean is 12 . 10^{-4} amp./km.^{2}. These quantities are of the same order as those deduced (for a slightly lower latitude, 39° N. instead of about 53° N., as in this paper) by Dr. Bauer from Miss van Vleuten's work; but fig. 2, in which curve (*b*) represents the results of Table II, shows that the two inferred current-variations are of nearly opposed types; the present results show a downward current-density (reckoned from the daily mean as datum) during the hours 9 to 13, whereas Dr. Bauer's indicate an upward current during this period.

The physical significance of Table II can be estimated only after a careful examination of the errors to which the results are subject. In this connection it is worthy of note that a magnetic force of 1 γ along the boundary, in the same sense all round, would correspond to a mean current density over the enclosed area of 5·5 . 10^{-3} amp./km.^{2}. Thus the maximum value of the line-integral, at 11 h., corresponds to an average force of 0·75 γ all round the boundary, and the *mean* deviation of the current-density from the daily mean corresponds to an average force of 0·22 γ along the boundary. The extent to which an error of 1 γ in ΔN or ΔW at each of the five stations would by itself affect the results of Table II is as follows, expressed in the same unit (10^{-4} amp./km.^{2}) as there:

Table III.

—	Greenwich.		Eskdalemuir.		Rude Skov.		Seddin.		de Bilt.		Mean.	
Error of 1γ in	ΔN	ΔW	ΔN	ΔW	ΔN	ΔW	ΔN	ΔW	ΔN	ΔW	ΔN	ΔW
Error in Δi	4	6	5	8	3	10	4	5	1	9	3·4	7·6

It is clear that the results depend much more on the accuracy of ΔW (or ΔD) than of ΔH or ΔN.

The estimates of Δi are affected by errors of two distinct kinds, which will be considered in turn. The errors of the first kind are due to inaccuracies in the data for the five stations, as representing the values of ΔH at those points. The errors of the second kind arise in the treatment of these data, which are made the basis of inferences as to the value of ΔH at intervening points.

The Errors in the Data used.

8. The errors in the data afforded by the observatories may be divided into (*a*) purely accidental errors arising in the measurement of the magnetographs, (*b*) errors of scale value of the magnetographs, and (*c*) changes of base-line value in the course of a day (any average error affecting the whole day is removed when the daily mean value of **H** is subtracted from the hourly values) due to temperature or other causes.

(*a*) The scale values of the magnetographs are indicated in Table I to the nearest unit in γ per mm. In general, the scale is 5 γ or less per mm., though in one case it is 10 γ. If the probable error of measurement is $0 \cdot 2$ mm., this corresponds to about $1 \cdot 0\ \gamma$. Each hourly value used is derived from 20 days, so that the hourly mean difference from the daily mean will have a probable error on this account of about $0 \cdot 2\ \gamma$. Ten such hourly means (two for each of five stations) are used in determining each hourly result given in Table II, and though they do not all equally affect the estimate of Δi (*cf.* Table III), the probable error in Δi due to this cause will not exceed $\frac{1}{3}$ unit in Table II. In some of the 25 line-integrals computed the error might rise to 2 or 3 units (10^{-4} amp./km.^{2}) in Table II, but scarcely to more. Consequently, errors of this kind are of little importance in our work.

Appreciable errors of measurement may also occur at some hours owing to slight errors in time-marking, causing the ordinate to be measured at a point on the curve not corresponding exactly to the Greenwich hour. For about three hours shortly before noon ΔW varies at the rate of nearly 15 γ per hour, and an error of $0 \cdot 5\ \gamma$ at these hours might arise in this way. But such errors would be unsystematic and, like those of measurement of the ordinates, without serious effect on the estimates of Δi.

(*b*) The scale values are not likely to be in error by so much as 2 per cent.; an error of this magnitude would affect ΔN and ΔW systematically on all the 20 days by an amount which would be greatest at the hours of maximum or minimum (10 h., 11 h., and 19 h. for ΔN, and 7 h., 8 h., 12 h., 13 h. for ΔW). At these times ΔN or ΔW would be in error by about $0 \cdot 5\ \gamma$: if such an error occurred at one station and in one element only, Δi at the corresponding hour would be affected by (on the average) 3 of the units of Table II, though the amount (*cf.* § 7, Table III) might rise to 5 units. Correspondingly larger errors would arise through inaccuracies in the scale values of more than one observatory and element, but it is unlikely that such inaccuracies would all affect Δi in the same sense. We may conclude that, at the hours named, errors of scale value may account for a few units of Δi in Table II, but that the contribution to the average deviation (12 units) is not likely to exceed 2 or 3 units.

(*c*) At Greenwich and de Bilt the hourly values of ΔN or ΔH are corrected for changes of temperature on the rare occasions when this is considered necessary (the daily temperature range is normally less than 0°·05 C. at de Bilt, and " too small to be measured with an ordinary thermometer " at Greenwich). In view of the small temperature ranges and temperature coefficients of the instruments, and the slight influence of ΔH at de Bilt on Δi (*cf.* Table III), any inaccuracy in the corrections will produce errors in Δi which, though likely to be systematic—that is, much the same from day to day at the same hours—will hardly amount to 1 unit of Table III. At Seddin the temperature of the room may vary by 2 degrees in a day, but the range inside the instrument cases is stated to be probably less than 0°·1 C.; in view of the very small temperature coefficients at Seddin, no appreciable errors seem likely to arise in Δi due to the Seddin temperature changes. At Rude Skov, Eskdalemuir and Stonyhurst the instrument room is underground, and the daily variation of temperature appears to be small, perhaps less than 0°·05 C. in general; no temperature corrections are applied. The error in ΔH or ΔN due to this cause is unlikely to exceed 0·3 γ at any hour. It would seem, on the whole, that errors in Δi due to temperature changes at the five observatories are likely to be small—probably less than 1 unit in the mean deviation.

The Errors due to Interpolation.

9. In the computation of $\int \Delta \mathrm{H} \,.\, \mathbf{ds}$ it was assumed that ΔH varies uniformly along each of the five arcs composing the contour. Light is thrown on the validity of this assumption by comparing the values of ΔN and ΔW for de Bilt with those obtained by interpolation between Greenwich and Seddin for the point nearest to de Bilt on the great circle passing through these two stations; this point is only 8′ distant (in latitude) from de Bilt. An approximate comparison showed considerable discrepancies both in ΔN and ΔW at certain hours. A more exact comparison was therefore made, as follows: assuming that ΔN and ΔW vary as linear functions of the latitude and longitude over the region considered, their values at de Bilt for each hour were obtained (i) by interpolation between Greenwich, Seddin and Eskdalemuir, and (ii) by interpolation between Greenwich, Seddin and Rude Skov. The linear relation leads,

in fact, to the following formulæ in the two cases: Δ refers to either ΔN or ΔW throughout, and the suffix indicates the observatory:

$$\text{(i)}\ \Delta_{dB} = 0\cdot507\ \ \Delta_G + 0\cdot417\ \ \Delta_S + 0\cdot076\ \ \Delta_E,$$

$$\text{(ii)}\ \Delta_{dB} = 0\cdot598\ \ \Delta_G + 0\cdot316\ \ \Delta_S + 0\cdot086\ \ \Delta_R.$$

The differences between the values of ΔN and ΔW obtained from these two distinct formulæ proved to be very small, the numerical average over the 24 hours being 0·16 γ for ΔN and 0·25 γ for ΔW. But these interpolated values differ much more from the actual values at de Bilt, as is shown by Table IV, which gives the actual values and the computed values (i) and (ii); for example, at 10 h. both ΔN and ΔW show discrepancies of about 3 γ.

Table IV.

Hour.	ΔN, in γ.			ΔW, in γ.		
	Computed.		Observed.	Computed.		Observed.
	(i)	(ii)		(i)	(ii)	
0	6·2	6·3	6·9	− 2·2	− 2·6	− 2·6
1	5·4	5·4	6·5	− 2·8	− 3·2	− 2·8
2	4·8	4·7	5·9	− 3·6	− 3·8	− 4·4
3	5·1	5·0	5·8	− 4·9	− 5·1	− 4·8
4	6·4	6·4	7·5	− 8·8	− 8·9	− 9·7
5	6·4	6·3	7·6	−13·5	−13·8	−15·4
6	3·4	3·1	3·8	−18·8	−19·0	−20·3
7	− 2·2	− 2·6	− 2·0	−21·7	−21·8	−24·0
8	− 9·1	− 9·4	−10·5	−21·5	−21·6	−22·7
9	−15·8	−16·1	−19·1	−15·2	−15·2	−15·2
10	−20·4	−20·5	−23·1	− 4·3	− 4·2	− 0·4
11	−20·7	−21·1	−22·9	10·9	11·2	16·4
12	−18·5	−18·7	−19·1	22·2	22·2	27·8
13	−14·1	−14·0	−14·6	25·7	26·0	28·5
14	− 8·1	− 8·0	− 7·6	23·2	23·4	22·1
15	− 2·9	− 2·7	− 2·6	16·4	16·9	14·7
16	1·9	2·2	2·1	9·8	10·4	7·5
17	6·5	6·7	7·6	4·9	5·4	2·6
18	10·5	10·7	11·3	2·1	2·6	1·0
19	12·2	12·4	12·7	1·1	1·2	1·2
20	11·4	11·6	11·3	0·4	0·4	0·7
21	10·6	10·6	10·2	0·2	0·1	0·7
22	10·0	10·1	11·0	− 0·3	− 0·4	− 0·4
23	9·3	9·4	9·3	− 0·8	− 1·1	− 0·8
24	8·7	8·9	9·3	− 1·1	− 1·5	− 1·7
Numerical Mean	9·35	9·44	10·14	9·76	9·98	10·24

10. In view of these differences it was decided to compute line-integrals afresh for a new and simpler contour approximately coinciding with the original

one, but composed of two meridian arcs ($\lambda = 1° 36'$ W. and $12° 44'$ E.) and two parallels of latitude ($51° 57'$ and $55° 35'$ N.), and to make interpolations in three different ways, by considering (1) only the four observatories of Greenwich, Eskdalemuir, Rude Skov and Seddin; (2) these and also de Bilt; (3) all six observatories. The differences made by the inclusion of data from the one or two additional observatories near the middle of two sides of the boundary should throw valuable light on the reliability of the results.

The meridianal sides of the new boundary curve are of length 404 km., and the sides along the northern and southern limits of latitude are of length 901 and 982 km.; the area enclosed is $0{\cdot}00937\ R^2$, or 380,000 km.^{2}. Along the meridianal sides only ΔN contributes to the line-integral, and only ΔW along the other sides.

In the first of the new computations ΔN or ΔW was taken to be uniform along each side, and as having the value obtained for the middle point of the side by linear interpolation between the four observatories, giving double weight to the two nearer ones (which is equivalent to taking the mean of formulæ (i), (ii) in the de Bilt calculation). The values of Δi obtained in this way are given in the second column (*c*) of Table V, and in curve (*c*), fig. 2; at some hours they show marked differences from the values obtained in the first calculation (Table II and curve (*b*)).

11. In the second new computation the information afforded by de Bilt as to the variation of ΔW along the southern arc of the boundary was used. De Bilt is $23'$ to the west and $9'$ to the north of the middle point P of this arc. By means of the two linear relations on which formulæ (i), (ii) of § 9 are based the differences were estimated, at each hour, between the values of ΔW at de Bilt and at P, due to these slight differences of latitude and longitude. The difference of latitude was found to be almost immaterial, having a maximum influence of $0{\cdot}3\ \gamma$. The difference of longitude had but little greater influence, the average for the 25 hours being about $0{\cdot}3\ \gamma$ and the maximum $0{\cdot}5\ \gamma$. These corrections were applied in order to reduce the de Bilt values of ΔW to those for P, though the changes scarcely affected the line-integral.

The resulting values of ΔW at P were different from the values obtained without considering de Bilt, indicating that the variation of ΔW along the arc is not linear. Assuming the variation to be represented by a graph which is of the next simplest form, namely, a parabolic arc instead of a straight line, the most probable mean value of ΔW along this side of the boundary is obtained by adding to the former value (regarded as the mean between the values at the two ends) two-thirds of the difference between this mean and the

Table V.

Hour.	*i*.		
	c.	*d*.	*e*.
0	23	22	19
1	12	5	1
2	− 8	− 1	− 4
3	−11	−12	−17
4	−13	− 2	−10
5	−20	0	−13
6	−15	2	−13
7	−21	10	− 5
8	−15	0	−13
9	− 9	− 9	−17
10	28	−24	−19
11	8	−63	−50
12	25	−51	−22
13	− 4	−37	− 9
14	− 5	15	32
15	−23	10	22
16	−32	11	12
17	−28	12	19
18	−13	8	0
19	15	15	14
20	9	5	1
21	25	16	8
22	25	24	25
23	28	23	18
24	21	20	16
Numerical Mean	17·2	15·6	15·0

actual value at P, inferred from de Bilt. This modified the contribution of this arc of the boundary to the value of the line-integral, and led to the new estimates of Δi contained in the third column (*d*) of Table V. Many of the values in this column differ appreciably from the corresponding ones in Table II; this is well shown in fig. 2 (curves *b*, *d*).

In a similar manner the Stonyhurst values of ΔH and ΔD were used to correct the mean value of ΔN along the eastern arc of the boundary curve, and further corrected values of Δi, given in the fourth column (*e*) of Table V and in curves of fig. 2, were thus obtained.

We can now estimate how far the necessity for interpolation between the stations for which data are available is likely to affect the reliability of our line-integral variations. The mean ordinate, without regard to sign, has the following values for curves *b*, *c*, *d*, *e*:

12, 17, 16, 15.

Curve *c* is based on four stations only, and may be expected to be the least accurate; curve *e* should be the most accurate, while curves *b* and *d* should be fairly similar. At first sight it appears from fig. 2 that curves *b*, *d*, *e* are moderately accordant, though they differ considerably from curve (*c*) in which the de Bilt data are not considered. But the mean numerical differences between the various curves are comparable with the above numerical mean ordinates: they are as follows:

b—*d*	*b*—*c*	*d*—*c*	*e*—*c*	*d*—*e*	*b*—*e*
7	14	21	18	9	9

As fig. 2 shows, the differences between curve (*c*) and the other three are comparable with the ordinates of the various curves; the mean differences between curves *b*, *d*, *c* are rather more than half the mean ordinates of these curves. While the change in the curves due to the inclusion of the Stonyhurst data (affecting one of the short sides of the boundary) is not small, it is far outweighed by the difference made according as the de Bilt data (affecting the integral along one of the long sides) are included or not. It is of interest to note that the corrections to curve (*c*) made by the inclusion of the de Bilt data (to give curve (*d*)) are, in 22 out of 24 cases (in the twenty-fifth case, for 19 h., there is no change), in the right direction to reduce the integral to zero, though the correction sometimes fails to achieve this, and in other instances overshoots the mark. It may be that if data were available for a further station midway along the other (northern) long side of the boundary, the differential effect of the corrections due to this and de Bilt would considerably reduce the mean ordinate. In any case it is clear that in the absence of such a station no great reliance can be placed on any of the curves (*b*) to (*e*), since the mean ordinates are comparable with the uncertainties in them. Since the uncertainty in the calculated earth-air current-densities is comparable with their magnitudes as thus estimated, the fact that both the estimates and the possible errors far exceed the direct measures of the diurnal variation of the current-density need cause no surprise; it implies merely that the available magnetic data are inadequate to determine the real current-densities. On the other hand they do suffice to show that at most only a very small fraction of the diurnal magnetic variations can be supposed not to possess a potential, since the errors in ΔN and ΔW, which would account for the non-zero line-integrals, are of the order 1 γ, whereas ΔN and ΔW themselves (*cf.* Table IV) range up to nearly 30 γ.

It does not seem possible to reduce the errors involved in the preceding calculations by any other choice of data at present available, but the evidence for the existence of a potential for the diurnal magnetic variations appears to us adequate. The onus of proof to the contrary seems to rest with any who may controvert the conclusion, and the most direct evidence that could be brought forward would consist of calculations of line-integrals such as are here given, but based on more accurate data: these would require to be specially obtained for the purpose. The undertaking would be difficult and costly unless new and simpler recording devices than those now in use should become available.

A New Interference Method of measuring the Surface Area of Film Catalysts. Part I.—*The Theory.*

By F. Hurn Constable, M.A., Ph.D. (London, Cambridge), Fellow of St. John's College, Cambridge.

(Communicated by T. M. Lowry, F.R.S.—Received March 7, 1928.)

One of the great difficulties presented in all investigations on the catalysis of heterogeneous gas reactions is the experimental preparation of a catalyst of definite mass associated with a measurable and reproducible surface area. Masses of metallic oxide, when reduced at low temperatures, give a metal which has often a considerable activity, but an entirely unknown surface area, which is liable to considerable variation. The use of supported catalysts is a considerable advance. The mass of the catalyst is fixed, and since the whole mass may be rapidly oxidised and reduced, leaving no massive metal beneath unattacked, the variations in the surface area presented to the gas phase is considerably restricted. The external surface will vary with the method of production of the film, and it is the purpose of this paper to describe a general method of measuring this surface, the determination of the surface area of these films being of vital importance in the study of their catalytic properties.

On the Meaning which may be assigned to the Surface Area of a Catalyst.

The area of a continuous bounded surface which is perfectly plane does not vary with the means of measurement adopted. When, however, irregularities are present in the surface, the surface may be considerably increased, the increase varying both with the magnitude of the irregularities and the means of estimation adopted. By some methods of measurement the irregularities are disregarded, by others they are included.

The measurement of length involves the same problem. A vessel sailing round the coast of England keeping within a few miles of the shore would travel a much smaller distance than would a man walking round the edge of the cliffs, and the man would travel less far than an insect making the same tour.

The smallest area that could be ascribed to the film is that of the inert support; the largest value for chemical purposes is the area of a surface composed of hydrogen atoms, all in contact with each other, and with the catalyst surface, which is thus completely covered.[*] There is grave doubt whether this maximum value could be obtained experimentally, for not only is there a pronounced lack of heterogeneity in the adsorbing power of a reduced metal for hydrogen, but the problem is also complicated by the solution of the gas in the metal.

The evidence shows that adsorption methods can only hope to measure a fraction of the true surface.

To make certain that the surface is covered with a homogeneous film one is driven to chemical methods of attacking the surface, since then it may sometimes be asserted that the whole surface is covered with a homogeneous film. If the mass of this film be measured ($w = \rho \,.\, x \,.\, S$), and the thickness x be independently estimated, then the value of S can be obtained if the density of the film be known.

When a gas reacts with a solid surface, and the specific volume of the product is greater than that of the original metal, then the surface becomes wholly covered with the reaction product, and, further, the rate of reaction is chiefly regulated by the rate of diffusion of the gaseous reactant through the film already present. These conditions lead to the production of very uniform films over a metallic surface.

A considerable amount of experimental investigation[†] has been made on the oxidation and reduction of metallic copper. During oxidation the surface is covered with a continuous film of oxide, which under some conditions shows interference colours. Direct spectrophotometric measurements of the intensity and wave-length of the light reflected from oxidised surfaces of iron, nickel and copper leave no doubt that interference is the main cause of production of the colour sequences;[*] and thus the colour of these films uniquely determines their thickness when the order of the colour and the refractive index of the film is known. We have therefore the means of developing a general method of evaluating the surface area.

---

[*] It seems possible that some method of electrodeposition of hydrogen from solution might be adapted to measure this area, but the difficulty is to obtain satisfactory evidence that the quantity of electricity passed has actually caused the production of a complete and unimolecular film of hydrogen over the surface.

[†] Joannis, 'C. R.,' vol. 158, p. 1801 (1914); Berger, *ibid.*, pp. 1502, 1798; Stahl, 'J. Soc. Chem. Ind.,' vol. 23, p. 1158 (1914); Tammann, 'Z. Anorg. Chem.,' vol. 111, p. 78 (1920); Tammann and Koster, *ibid.*, vol. 123, p. 196 (1922); Hinshelwood, 'Roy. Soc. Proc.,' A, vol. 102, p. 318 (1922); Palmer, 'Roy. Soc. Proc.,' A, vol. 103, p. 444 (1923); Evans, 'J. Chem. Soc.,' vol. 127, p. 2484 (1925), and 'Roy. Soc. Proc.,' A, vol. 107, p. 228 (1925); Dunn, 'Roy. Soc. Proc.,' A, vol. 111, p. 210 (1926); Pilling and Bedworth, 'J. Inst. Metals,' vol. 29, p. 529 (1923); Constable, 'Roy. Soc. Proc.,' vol. 107, p. 270 (1925), vol. 115,

Methods of using Interference Films to Determine Surface Area of Metallic Film Catalysts.

Let S be the surface area of metal of mass M, forming W grams of compound when completely converted by the gaseous reagent.

w be the mass of the film covering the metal at time t.

q be the thickness of a film of air showing the same colour.

μ be the refractive index of the oxide for wave-length λ.

C be the electrical conductivity of the metallic film when partially attacked, and

C_0 be the initial conductivity.

λ be the wave-length of the absorption or reflection maximum in the spectrum of the reflected light.

I.—Direct measurement of S from measurements of w and spectrophotometric observations on the light reflected from the coloured film.

The thickness of the film $x = (2n - 1)\lambda/4\mu$; $n = 1, 2, 3, \ldots$, etc., and λ is the wave-length of the absorption band, or $x = n\lambda/2\mu$; $n = 1, 2, 3, \ldots$, when there is a maximum of reflection.

Hence $w = S \cdot \rho \cdot (2n - 1)/4\mu$ or $S\rho n/2\mu$; thus, on plotting w against $n'\lambda/\mu$,[†] a straight line is obtained from which the area S is determined. S/M is found by a separate determination of M.

---

p. 570 (1927), vol. 117, p. 376 (1928), and 'J. Chem. Soc.,' vol. 127, p. 1576 (1927), and 'Nature,' vol. 118, p. 730 (1926), vol. 119, p. 349 (1927), vol. 120, p. 769 (1927). The fact that reduction of the oxide only occurs at the metal oxide interface is well brought out by Pease and Taylor, 'J. Amer. Chem. Soc.,' vol. 43, p. 2179 (1922); Jones and Taylor, 'J. Phys. Chem.,' vol. 27, p. 623 (1923).

[*] 'Roy. Soc. Proc.,' A, vol. 117, p. 376 (1928), and vol. 115, p. 570 (1927).

[†] n' represents the required integer.

II.—From data on the spectrophotometric analysis of the reflected light and the conductivity of the film.

The fall in conductivity is taken as a measure of the amount of the metallic film that is converted into the covering film, therefore

$$\frac{C_0 - C}{C} = \frac{w \,.\, p}{M} = \frac{S\rho p\lambda}{M} \,.\, \frac{2n-1}{4\mu} \quad \text{or} \quad \frac{S\rho p\lambda}{M} \,.\, \frac{n}{2\mu},$$

where p is the percentage of metal in the covering film. Hence, writing Δc for the fall in conductivity we have

$$\frac{S}{M} = \frac{\Delta c \,.\, p}{\rho \,.\, C} \,.\, \frac{(2n-1)\,\lambda}{4\mu}$$

for the absorption bands and

$$\frac{\Delta c \,.\, p}{\rho C} \,.\, \frac{n\lambda}{2\mu}.$$

for the reflection bands. It is convenient to plot $\Delta c/C$ against $n'\lambda/\mu$, the slope of the line determining S/M.

III.—From observation of the surface colour expressed in terms of the equivalent air thickness of the covering film and the mass of the film.

The mass of the covering film may be expressed as

$$w = S \,.\, \rho \,.\, q/\mu.$$

Thus $S = \mu \,.\, w/\rho \,.\, q$, and if a separate determination of M be made, the value S/M is determined.

IV.—From simultaneous observations of the surface colour and the electrical conductivity of the metallic film.

We have

$$\Delta c/C = w \,.\, p/M,^*$$

whence

$$\Delta c/C = p \,.\, \rho \,.\, q \,.\, S/\mu M,$$

or

$$S/M = \Delta c\mu/C \,.\, q \,.\, \rho \,.\, p.$$

It is convenient to plot q/μ against $\Delta c/C\rho p$, the straight line slope enabling S/M to be measured.

---

* It is to be observed that p, being the percentage of metal in the metallic oxide, is altered only slightly by uncertainty in the nature of the oxide formed. Thus, if the composition of the surface film is only approximately known, the method is still tolerably accurate. The refractive index can be found by reflecting polarised light from the film.

V.—From the initial values of the oxidation constants determined from observations of the surface colours, together with the mass of oxide produced, or the change in conductivity of the metallic film.

Pilling and Bedworth, and Dunn (*loc. cit.*), have found that $w^2 = kt$, where w is the total mass of oxide formed, is an approximation to the law of formation of the film obeyed with fair accuracy by the sintered metal but showing very distinct fall in the value of the constant k for the activated metal. The law is useful in enabling the initial value of k to be extrapolated.

Thus

$$2w\frac{dw}{dt} = k_w$$

hence

$$\frac{dw}{dt} = \frac{k_w}{2w} = \tfrac{1}{2}\sqrt{\frac{k_w}{t}}.$$

Similarly

$$\frac{dq}{dt} = \tfrac{1}{2}\sqrt{\frac{k_q}{t}}; \quad \frac{dx}{dt} = \tfrac{1}{2}\sqrt{\frac{k_x}{t}},$$

and

$$2(C_0 - C)\frac{dC}{dt} = k_c,$$

therefore

$$\frac{dC}{dt} = \tfrac{1}{2}\sqrt{\frac{k_c}{t}}.$$

The rate of formation of the layer as a fraction of the total amount of oxidised film possible

$$\frac{e_M}{e_0}\frac{\rho S}{\mu M}\frac{dq}{dt} = \frac{1}{C_0}\frac{dC}{dt}$$

is equal to the rate of disappearance of the metal expressed similarly. Thus

$$\frac{S}{M} = \frac{e_0}{e_M} \cdot \frac{\mu}{\rho C_0}\sqrt{\frac{k_c}{k_q}},$$

where e_0 and e_M are the equivalent weights of the metal and oxide respectively.

If w be measured the relation

$$\frac{\rho \, . \, S \, . \, dq}{\mu \, dt} = \frac{dw}{dt}$$

becomes

$$S = \frac{\mu}{\rho}\sqrt{\frac{k_w}{k_q}},$$

and an outside measurement of M has to be made.

VI.—Comparative values of S_1/S_2 may be obtained when the nature and the refractive index of the film are uncertain.

For

$$\frac{\left(\frac{dq}{dt}\right)_1 . S_1}{\left(\frac{dq}{dt}\right)_2 . S_2} = \frac{\left(\frac{dC}{dt}\right)_1}{\left(\frac{dC}{dt}\right)_2},$$

whence

$$\frac{S_1}{S_2} = \frac{\left(\frac{dC}{dt}\right)_1 . \left(\frac{dq}{dt}\right)_2}{\left(\frac{dC}{dt}\right)_2 . \left(\frac{dq}{dt}\right)_1},$$

and no value of p or μ is assumed.

Reasons for the Introduction of the New Method.

The hot-wire technique, introduced by Hinshelwood, enables gas reactions proceeding at high temperatures to be studied in a closed vessel surrounded with melting ice. Pressure measurements are sufficient to describe the changes taking place in the vessel. Many useful data have been obtained by Hinshelwood and his co-workers by the use of this method.

The nature of the surface presented by the wire is a very uncertain factor, and so the wire has been replaced by a continuous, but very thin metallic film supported by china clay. This film has been obtained by electrodeposition by a method which is described in the following paper, and its mass is accurately known. It may be heated to any desired temperature by the passage of an electric current through the film, and the treatment to which the surface is subjected can be varied at will. At the same time the methods previously described enable the visible surface to be measured.

The catalyst is therefore standardised in a way which has not been possible previously. It is the purpose of this series of papers to apply the method of measurement to a whole series of metallic films.

A New Interference Method of Measuring the Surface Area of Film Catalysts.—Part II. Nickel—A Method of Preparation of the Film, and an Apparatus for Activation, and Study of the Surface Area.

By F. H. Constable, M.A., Ph.D. (London, Cambridge), Fellow of St. John's College, Cambridge.

(Communicated by T. M. Lowry, F.R.S.—Received March 7, 1928.)

Trials were made in order to obtain a reliable method of producing thin films of metals on china clay, and a technique for their activation, and study of the surface area. The uniformity of the mean thickness of the film is essential, as the final intention was to use these same films for the study of catalytic behaviour in the same manner as Hinshelwood has used metallic wires. The films finally obtained were stable to a number of oxidations and reductions, but if these were very prolonged, collapse of the plane structure of the film would take place, due to the curling of small portions of the film, giving the whole a fluffy white appearance. The surface area was thus further increased after lapse of a considerable time.

The electrical conductivity of the film was measured at frequent intervals during the oxidation, and spectrophotometric and visual observations of the surface colour were made: thus the area was easily measured. A special quartz apparatus was used for the activation, sintering and measurement of the surface area of the nickel.

Preparation of Active Nickel Films.

The application of finely ground nickel oxide, made into a stiff paste with oleic acid, to the surface of a china clay rod resulted in a film which was stable to very gentle ignition.* When reduced the film was grey, metallic and continuous to the eye, but the resistance of 8 cms. of the rod exceeded 10^6 ohms, showing that the film was composed of discrete particles of nickel completely separated from each other.† This film would show bright interference colours on oxidation, but owing to the lack of conductivity no useful information on the surface area could be obtained. The former method of producing films by means of the oxide paste with oleic acid was therefore abandoned in favour of an electrical method of deposition of the metal on graphite-coated china clay rods.

---

* *Cf.* Constable, 'Roy. Soc. Proc.,' A, vol. 117, p. 376 (1928).

† This was confirmed by the microscope.

The Preparation of Uniform Graphite Films on China Clay.

The china clay rods were first rubbed with purified powdered graphite, slightly damped with oleic acid. A black uniform layer covered the china clay rod, which was then polished with an old and well-washed linen rag coated with dry graphite flakes. In this state the film does not conduct electricity. On slightly warming, however, a slight conductivity is developed owing to the fact that some of the oleic acid is driven off from between the graphite plates. Ignition in a flame burns off the graphite, so an electrical method was adopted, the final driving off of the oleic acid being accomplished by the heating of the film itself by a steady current passing through it. A variable resistance of 200 ohms was placed in series with the graphite-coated rod, and a potential difference of 110 volts applied. Initially the current was of the order of 0·05 amp. After a few seconds the film becomes heated, and the oleic acid is vaporised, the resistance falling rapidly at the same time. Resistance is placed in the circuit during the removal of the oleic acid, and continuously increased so that the current might be maintained constant at 0·20 amp. It is important that the film should not be overheated in the initial stages of the vaporisation, otherwise portions of the film may be raised to redness and the uniformity destroyed. Finally, a current of 0·40 amp. was allowed to pass for about twenty seconds. While still hot the rod was rubbed with a similarly coated rod, and continuously rotated, so that the whole film is hardened and pushed down into the china clay support.

The film presents a uniform polished surface which is jet black in colour. The resistance of 8 cms. length of the rod was about 200 ohms. Wetting the dried film with oleic acid was not sufficient to cause any appreciable change in the conductivity, showing that the high-temperature sintering had formed a continuous film of graphite, which the oily oleic acid could not separate again. It is advisable to proceed with the metal plating of the graphite at once, when very good metallic deposits can be obtained. If the film is old, the graphite base is apt to flake off slightly and spoil the quality of the deposit.

The Electrodeposition of Nickel on Graphite.

Carbon is soluble in molten metal, and the amount dissolved increases with the temperature till 2100° C. is reached. A carbide of nickel of the formula Ni_3C seems to be formed, which dissociates at lower temperatures. When the temperature is as high as 1600° C., the dissociation is extremely rapid, so that at low temperatures the amount of carbon in combination is negligible. Thus the supporting graphite can be neglected in experimental work on catalysis by, and the oxidation of, films of nickel deposited thereon. The graphite-coated rod was wetted with distilled water till saturated, so that a large quantity of the nickel solution from the plating bath should not be taken up by the porous china clay, and give rise to various difficulties when ignited. The following baths were used[*] :—

Bath I.—Nickel ammonium sulphate with ammonia

6 grams nickel ammonium sulphate.
1 gram nickel chloride.
2·5 c.c. 10 per cent. ammonia solution.
100 c.c. water.

Bath II.—Nickel ammonium sulphate with boric acid.

6 grams nickel ammonium sulphate.
1·5 grams ammonium chloride.
1·5 grams boric acid.
100 c.c. water.

Bath III.—Nickel sulphate with boric acid.

70 grams nickel sulphate.
3 grams sodium chloride.
6 grams boric acid.
500 c.c. water.

The current efficiency in deposition from these three baths was studied. The graphited china clay rods were weighed on a microbalance[†] after continued desiccation in a vacuum desiccator, the current passed for a series of times—10, 20 and 40 minutes—and the rod washed in running distilled water for an hour, then dried in a desiccator and re-weighed. At the same time a blank experiment was done. A rod was treated in exactly the same way, but no deposition of nickel was made thereon: the correction necessary because of the absorbed salts from the bath could be worked out. The electrochemical equivalent of nickel was taken as 0·000304 gram amp. per second.

---

[*] For discussion and theory of plating baths, see—

'Principles of Electrotyping and Electroforming.' Blum and Hogaboom. MacGraw Hill Book Company, New York and London (1924).

'Modern Electroplating.' W. E. Hughes. Oxford Technical Publications. Henry Frowde and Hodder and Stoughton, London (1923).

[†] I am indebted to Dr. W. H. Mills, F.R.S., for the use of this instrument.

Results.

	Mass of nickel deposited.	Mass equivalent to quantity of electricity.	Current efficiency.
Bath I—			
Experiment 1	2·03 × 10^{-3} grams	2·11 × 10^{-3} grams	96·3
,, 2	4·39	4·64	94·6
,, 3	11·74	12·03	95·6
Bath II—			
Experiment 1	1·74 × 10^{-3} grams	1·99 × 10^{-3} grams	87·5
,, 2	3·73	4·32	86·3
,, 3	10·31	11·95	86·1
Bath III—			
Experiment 1	2·36 × 10^{-3} grams	2·48 × 10^{-3} grams	95·0
,, 2	4·83	5·17	95·3
,, 3	12·28	12·96	94·9

The current density in all cases was not much different from 0·2 ampere per square decimetre. The deposits from Bath I were apt to be coloured yellow and to peel when thick; from Bath II quite white deposits were obtained, but again it was difficult to obtain considerable thickness; from Bath III very white deposits were obtained, and thickness of 0·50 mm. was easily obtained, and, moreover, the film could be burnished on the rod. In order to deposit a film of known mass it is only necessary to measure the current passing, and the time, for accurate measurements to be made. Moreover, since the mass to be measured is of the order of 10^{-3} grams, it is evident that the electrical method is to be preferred over the physical or chemical method.

Apparatus for insuring a Uniform Deposit.

Owing to the high resistance of the graphite films, the electrodeposition of nickel takes place first on the end nearest to the metal connection, and the white nickel deposit grows downward. Thus the upper end of the rod tends to accumulate a thicker deposit than the bottom, when the deposition occurs on a still rod. To produce uniform films an apparatus was used which lifted the graphite-coated rod out of the electrolyte as fast as the deposit grew downwards. The rod B was attached by a terminal screw in a stout copper arm to the float C, heavily weighted at the bottom, which could rise, guided by the tube E. Glycerine was used as lubricant. A table was made of the rate of flow through

D, and the corresponding rate of rise of water level on the scale F, which measures the rate of rise of the rod B on which the nickel is being deposited. A preliminary experiment was made using a current of 0·01 ampere and measuring the rate of growth of the nickel film downwards, 4·3 cms. in twenty minutes. During the deposition on a similar rod the deposit would be raised at this rate. The rod was finally replaced upside down and a final nickel coating given with the rod still. Readings of the current passing were plotted against the time and the quantity deduced from the area beneath the curve.

Since the rods were of porous china clay with a central hole through them, care was taken that the electrolyte from the bath was not appreciably retained by the rods. Before plating, the graphite-coated china clay rod was left in distilled water, so that it could absorb water to saturation, and after deposition the rod was left in running distilled water for four hours. The rod was dried gently by filter paper, then by evacuation in the furnace for one hour at 200° C. No appreciable oxidation of the nickel took place.

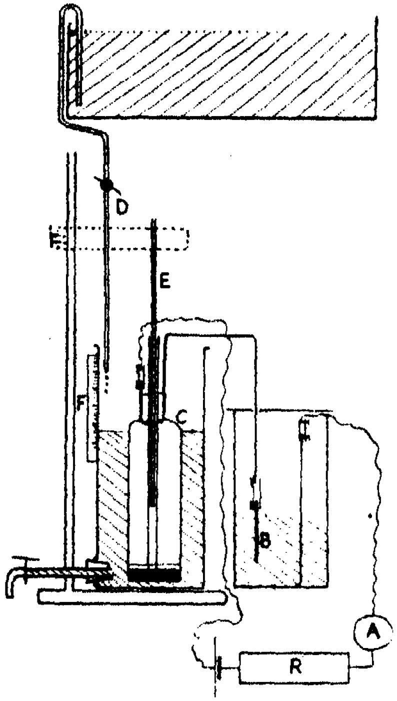


Fig. 1.—Apparatus for Electrodeposition of Nickel Films on Graphite.

Apparatus for the Activation, Sintering and Study of the Surface Area of Nickel Films.

The apparatus is shown in fig. 2. The pyrex glass tube previously used* was replaced by one of silica A, wound uniformly with bare nichrome wire. The furnace could be raised to 600° C. with ease by current at 220 volts, regulated by the rheostat R_1 and ammeter A_1. It is advisable, in practical working with bare heating wires, to renew the wires each month in order to avoid a breakdown during the oxidation run, and the consequent loss of a series of experiments. The temperature was measured by an iron-eureka thermocouple B in a silica sheath* placed centrally in the tube, the cool junction being in the thermos flask in ice, C. The couple, used with the millivoltmeter V_4 and the standard resistance coils R_4, had the advantage of possessing an almost linear calibration curve over the whole scale. For rapid evacuation the small silica tube was connected to the electrically driven double acting fleuss pump D. Two gas holders were provided, E and F, supplying carbon monoxide or hydrogen for the reduction of the nickel oxide to metal. The manometer K measured the pressure in the apparatus. The apparatus was rapid in working, some thirty oxidations and reductions being possible in five minutes if the temperature was elevated sufficiently, this being the chief factor regulating the speed of the process. When the furnace was very hot it was difficult, owing to the radiation from inside, to distinguish the colour of the film under observation. The pointolite G gave a sufficiently intense illumination, when focussed on the film, for the colour of the inside rod to be easily visible.

---

* 'Roy. Soc. Proc.,' A, vol. 115, p. 570 (1927).

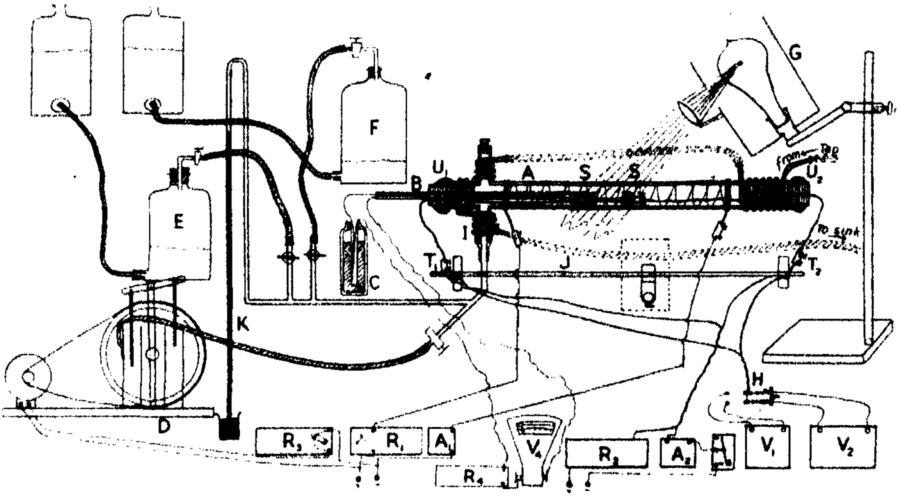


FIG. 2.—Apparatus for Oxidation and Reduction of Film, simultaneously observing Colour and Conductivity.

Method of measuring the Electrical Resistance of the Film.

The film, supported on china clay rods, was clamped between the two silver terminals S, and stout silver wires lead from the furnace to the terminals T_1 and T_2, which communicated, by way of R_2 and A_2, the momentary current necessary to measure the resistance by the voltmeter system connected to the change-over switch H. The silver wires passed through the rubber bungs stopping the ends of the furnace, and were sealed with Faraday wax. The stoppers were kept cool by the compo-pipe windings U_1 and U_2: incidentally, these served also to earth the ends of the silica tube and help to prevent leaks. The long glass rod J served to keep the silver wires well insulated till they joined the terminals T.

---

* I am much indebted to Col. C. T. Heycock, F.R.S., for this sheath and a supply of pure metals to calibrate the couple.

A rapid method of measuring resistance is essential. The first method tried was to use a galvanometer suitably shunted, and to pass 1/5000 ampere through the metallic film and note the deflection. When the temperature of the furnace became high, several factors, which had previously been of little importance, began to become serious. In spite of the earthed ends of the tube and the protection afforded by the insulating system, a persistent leak was discovered through the silica tube. The sign of the potential difference of the leak changed as the heating current was reversed. It became evident that the small current 1/5000 ampere did not give a sufficiently large potential difference to be measured with any accuracy under the given conditions. A large current

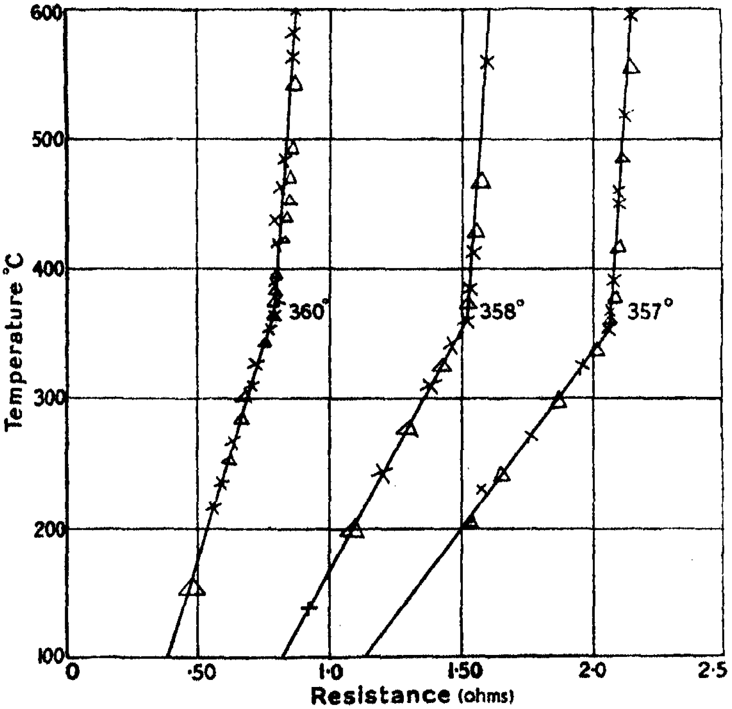


Fig. 3.—Showing existence of transition point about 360° C., independent of the presence of graphite and thickness of film.

of 1 ampere was accordingly passed through the film from the 110-volt mains, regulated by the rheostat R_2 and ammeter A_2. A heating effect was naturally produced, a rise of temperature of 66° C. being produced in 90 minutes. Thus it is necessary to restrict the time interval for which the current is left on very considerably. The potential difference was measured by a calibrated voltmeter, and a maximum depression of the tapping key of 3 seconds was the longest necessary to secure an accurate reading of the resistance. When thick films were studied, it was necessary to be able to measure smaller voltages of the order of 0·10 volt accurately, so a change-over switch H was provided for this purpose, and a millivoltmeter V_2, in addition to the calibrated voltmeter V_1, with three scales, 0–3, 0–6 and 0–30 volts.

The Effect of Temperature on the Film Resistance.

The temperature resistance curves were plotted for nickel films of various thicknesses on graphite-coated china clay rods. The magnitude of the error caused by the three seconds depression of the key is seen to be negligible, and a very interesting break in the temperature resistance curve was found at about 360° C., and this appeared independently of the thickness of the film and the presence of graphite. The transition temperature in the nickel system has been recorded.* It is interesting to notice that the lowest temperature at which nickel can be reduced (about 320° C.) is distinctly lower than the transition temperature, so that it is possible that the most active catalysts have a markedly dissimilar structure from the moderately active ones produced at higher temperature.

Determinations of the Area per Unit Mass of the Nickel Film, and Comparison with the Area of the Support.

Simultaneous measurements were made of the thickness of the surface film of oxide on the metallic nickel and the electrical conductivity of the film.

If C be the initial conductivity of the nickel film ;
p be the percentage of metal in the oxide ;
ρ be the density of the oxide ;
x be the actual thickness of the oxide film ;
t be the time ;

---

* Werner, 'Z. anorg. Chem.,' vol. 83, p. 275 (1913); Jäneke, 'Z. angew. Chem.,' vol. 31, p. 229 (1918); Baikon, 'Int. Z. Metallog.,' vol. 6, p. 115 (1914); Jäneke, 'Z. Electrochem.,' vol. 25, p. 9 (1919).

then the linear rate of growth of the oxide film is given by dx/dt, and if S be the surface area of the interface between the oxide and the metal, the quantity $p\,.\,\rho\,.\,S\frac{dx}{dt}\Big/M$ measures the rate of disappearance of the metal as a fraction of the total amount of metal present. The same rate of disappearance is also measured by $\frac{1}{C_0}\frac{dC}{dt}$. Thus,

$$\frac{p\,.\,\rho\,.\,S\,.\,\frac{dx}{dt}}{M}=\frac{1}{C}\,.\,\frac{dC}{dt}$$

or

$$\frac{S}{M}=\frac{1}{C_0\,.\,\rho\,.\,p}\,.\,\frac{dC}{dx}. \qquad (1)$$

Thus, the slope of the curve C, x curve, $\frac{dC}{dx}$ gives the area per unit mass of nickel. The majority of the curves obtained by plotting C and x were straight lines, but with the most active catalysts the slope sometimes showed a tendency to decrease as the oxidation proceeds. This was attributed to the decrease of the interfacial surface area on the thickening of the oxide film. The film was activated by oxidation and reduction at 580° C. five times. It was then reduced at the lowest possible temperature by hydrogen, and the curves shown in figs. 4 and 5 were obtained, each oxidation taking place at 330° C. The rapid slowing down of the initial rate of oxidation at these low temperatures is very marked.

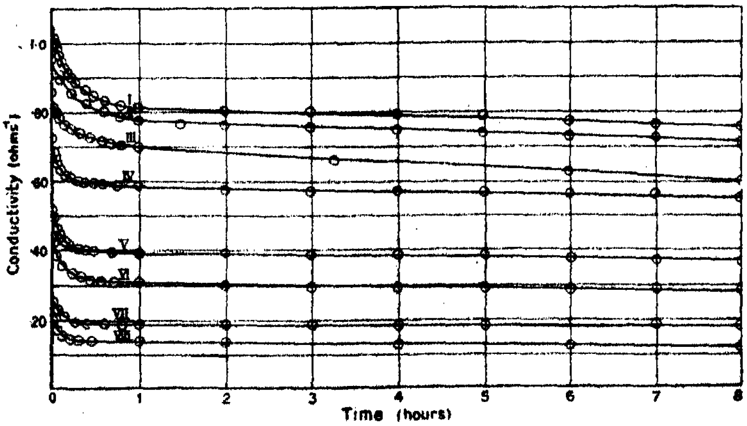


FIG. 4.—The Conductivity Time Curves 330° C.

The series of measurements was interrupted after curve IV, and a series of oxidations and reductions rapidly made at 563° C. The observations of the thickness of the oxide film and the conductivity are given in fig. 6. During the continued oxidations and reductions the initial conductivity persistently fell, owing to the continued opening of the structure of the film. The areas per gram of nickel are shown in Table II.

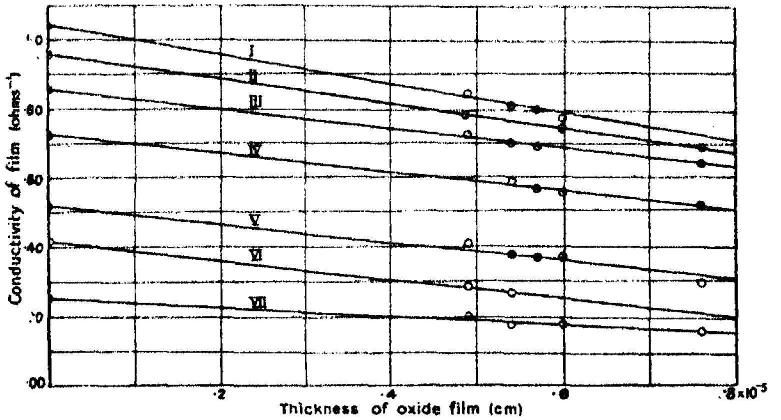


Fig. 5.—The Conductivity and Thickness of Oxide Film.

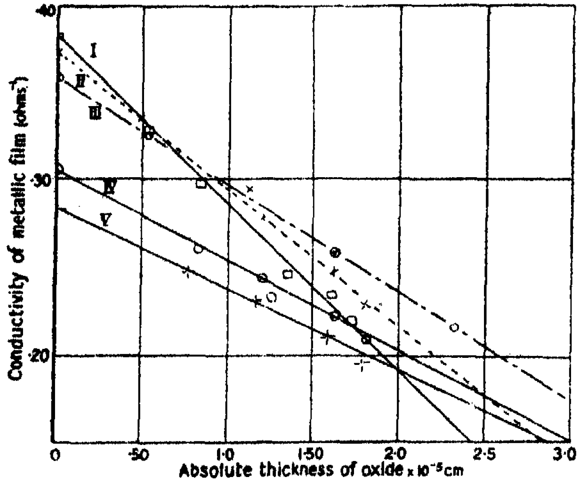


Fig. 6.—The Conductivity and Thickness of Oxide Film.

Table II.

No. of oxidation after reduction by hydrogen.	Temperature of		Area.	Ratio of actual area to area of support.
	Reduction.	Oxidation.		
	° C.	° C.	Sq. cms. per gm.	Times.
Figs. 4 and 5—				
Curve 1	350	350	$7·9 \times 10^{3}$	3·2
,, 2	330	330	6·5	2·5
,, 3	328	330	6·3	2·4
,, 4	328	330	6·4	2·5
Fig. 6—				
Curve 1	563	563	5·2	2·0
,, 2	563	563	4·2	1·62
,, 3	563	563	3·43	1·32
,, 4	563	563	3·47	1·36
,, 5	563	563	3·45	1·33
Figs. 4 and 5—				
Curve 5	328	330	7·0	2·7
,, 6	324	330	7·8	3·0
,, 7	324	330	8·5	3·3
,, 8	324	330	11·7	4·5

Discussion of Results.

The area on activation seems to pass through a minimum and then rises. The effect is accompanied by a fall in the conductivity of the film, and microscopic examination shows this to be due to cracking in the film, the strain set up by the alternate oxidation and reduction being sufficient to break up the film slowly if continued. After being activated, raising the temperature to 563° C., with alternate oxidations and reductions, demonstrated that two consecutive oxidations and reductions were necessary before the film would settle down to an equilibrium state. The terminal area was less than that shown by metal reduced at the lowest possible temperature.

Summary.

Metallic films made by the oleic acid method from the oxide on china clay failed to conduct electricity, so that it was necessary to introduce a new method. Graphite-coated china clay rods were used as the foundation for an electrolytic deposit. The quantity of nickel deposited corresponded to an efficiency of about 95 per cent. in the case of the nickel sulphate bath. The colour of the deposit varied widely with the bath used. By measuring the current and the time,

the mass of nickel deposited could be accurately ascertained. The activation and oxidation took place in a quartz tube, the apparatus being of special design. The maximum area of the film per unit mass of nickel was measured by the interference method for activation at the lowest possible temperature. The attainment of a terminal state after alternate oxidation and reduction was also studied at 563° C. The maximum increase in surface area found was of the order of five times, and was reduced to 1·3 times by reduction at 563° C.

The Nodal Cubic Surfaces and the Surfaces from which they are derived by Projection.

By G. Timms, St. John's College, Cambridge.

(Communicated by H. F. Baker, F.R.S.—Received January 17, 1928.)

The cubic surfaces have been classified according to the character of their singularities by Schläfli and by Cayley, who find that there are 21 types in addition to ruled surfaces. In their treatment of the matter each case is considered separately by algebraical methods, and there is a marked lack of any simple unifying principle, which it is the object of this paper to supply.

A means whereby this can be done is suggested by the theorem that every surface is the projection of a non-singular surface in higher space.† The considerations employed in the proof of this result are somewhat abstruse, and the purely geometrical significance is obscure, so that the more detailed examination of particular cases is of genuine interest. Accordingly, the subject of this paper is the generation of the various nodal cubic surfaces by the projection of non-singular surfaces, specifically the non-ruled surfaces of order n in space of n dimensions (denoted throughout by F^n); it will be shown that these arise by the projection of the same surface, F^9.

Incidentally the F^n which possess singularities are enumerated.

The various cubic surfaces are also generated by projecting the intersection of a suitable pencil of quadrics in four dimensions [§ 8*b*].

---

† B. Levi, 'Atti R. Acad. Sci. Torino,' vol. 33, p. 72 (1897).

I should like to thank Prof. H. F. Baker for his advice and encouragement during the preparation of this paper.

The symbol $[k]$ is used throughout for a flat space of k dimensions; the term solid is applied to a $[3]$, the term prime to an $[r-1]$ considered as lying in $[r]$.

§ 1. *Outline of the Method employed.*

1*a*. It is known [*v*. § 4*a*] that *for all* F^n, *i.e., non-ruled surfaces of order n in space of n dimensions,* we have $n \leqslant 9$; *and that all, save only a certain surface denoted by* F^{8*}—*or* F^8 (A^*)—*may be obtained by projecting* F^9 *from a* $[9-n-1]$, *or, which is the same thing, successively from* $9-n$ *points.* To preserve the equality of the order of the surface and the dimension of the space containing it, the sum of the multiplicities upon the surfaces of these $9-n$ points must be $9-n$, so that: *either* one of them is multiple, which case is shown in § 7*a* to lead only to *ruled surfaces and to the projections of the Veronese surface*; *or* the $9-n$ points are all simple points, *i.e., the* F^n*s so arising, for any given value of* n, *can be obtained from those for which n is one greater, by projecting these from all possible, essentially different, simple points thereon.* The result of this process is exhibited schematically in the tables appended to § 8.

Since the surfaces are, apart from their lines, homogeneous [§ 4*b*i], *the points are distinguished only according to the line or lines on which they lie.* These will be indicated symbolically by the Y in F^n (X) — (Y), *i.e.*, the surface F^n, with the singularities X, projected from the point Y. F^n (X) — (P) is used when the projection is from a general point of the surface.

It may be possible, in obtaining a given F^3, to project on to dimensions considerably less than [9] before a node appears [*v*. table in § 8].

1*b*. The criteria of distinctness for the F^n so obtained are the nature of the nodes, the lines, and the existence of their intersections. But no special account is taken of an F^3 on which three lines meet at a simple point, and therefore lie in the tangent plane, which is then said to be an osculatory plane. This circumstance cannot arise on an F^n, $n \geqslant 4$, for on projecting from general points of the surface into an F^4, it would be preserved. But the F^4 is the intersection of two quadrics [§ 4*b*i] which would therefore both contain the plane. This distinction will be neglected.

§ 2. *Description of the Singularities on a Cubic Surface, or other* F^n.

2*a*. Excluding the three types of cone and the two types of proper scroll [*v*. § 2*e*] there remain, according to the number and nature of their nodes, 21 types of cubic surface. These have been treated by Salmon,[†] Schläfli,[‡] who first distinguished the 21 types, Cayley,[§] and Bobek.[‖]

2*b*. No cubic surface, save a cone, has a point of multiplicity greater than two. The double points are classified[¶] as conic- (or cnic-) nodes C_2, binodes B, or unodes U, according as the nodal cone is a proper cone, two distinct planes (biplanes), or a repeated plane (uniplane).

A binode is denoted in general by B_3, but if the edge, *i.e.*, the intersection of the biplanes, lies on the surface, by B_4. If one of the biplanes touches the surface along the edge, the biplane is known as torsal, and the singularity denoted by B_5; if it meets the surface in the edge counted thrice, as oscular, and the singularity denoted by B_6.[††] Similarly to ordinary, torsal, and oscular uniplanes correspond the singularities U_6, U_7, U_8. In every case the suffix denotes the reduction in the class of the surface.

2*c*. This specification of these singularities is intelligible if applied to any surface in any number of dimensions, and the notation will be used for the F^n: it comprises all singularities occurring on the F^n except B_4^* [§ 5*f*] and a nodal line [§ 2*e*].

The (quadric) nodal cone is contained in a solid,[‡‡] the nodal solid, which is such that every prime through it meets the surface in a curve having a 4-ple point at the node. The latter property defines this solid at a unode.

2*d*. A line of a surface through a node thereof will be called a ray, one through two nodes, an axis.

[†] 'Camb. & Dub. Math. J.,' vol. 4, p. 256 (1849). Giving suitable multiplicities to account for the 27 lines and 45 tri-tangent planes.

[‡] 'Phil. Trans.,' vol. 153, p 193 (1863). Giving sub-classification according to reality of lines.

[§] 'Phil. Trans.,' vol. 159, p. 231 (1869); 'Coll. Pap.' VI, p. 359. Giving tangential equations.

[‖] 'Wien. Sitzungsb.,' vol. 96, p. 355 (1888). Generation by related pencils of planes and quadrics. Also enumeration based on the arrangement of the rays through any one node. Numerous other references in Pascal, 'Repert. d. höh. Math.,' vol. II_2, p. 824 (Teubner, 1922). See also Klein, 'Math. Ann.,' vol. 6, p. 551 (1873); 'Gesamm. math. Abh.,' vol. 2, p. 11, who obtains various cubic surfaces by continuous deformation.

[¶] Salmon, 'Analytical Geometry of 3 Dimensions' (5th ed., Longmans, 1915), vol. 2, § 522.

[††] The plane and the singularity are regarded as ordinary of the type, if the former touches or osculates the surface along a line on account of the presence of another node upon the surface; thus a biplane can be properly torsal only along the edge.

[‡‡] If, as throughout, $[k]$ denotes a flat k-dimensional space, a solid is a $[3]$, and a prime in $[r]$ is an $[r-1]$.

a, *b*, *c*, *d* invariably denote an axis, ray, edge and mere line respectively, though *d* is occasionally replaced by a purely numerical symbol.

To the self-explanatory symbols such as $F^n\,(B_3 + C_2)$ is added F^n (A) for a surface without nodes (Anautotomic).

2*e*. A proper cubic scroll (or ruled surface) contains a double line, with two tangent planes at each point of the line: these coincide for two points, known as the pinch-points.

Such surfaces, it will appear later [§§ 7*cii* and 7*ciii*] are projections of the Veronese surface. All surfaces, including ruled surfaces, of order n in space of n dimensions, which are projections of the Veronese surface, will be denoted by $\overline{F}^n$. In particular, the cubic scrolls will be denoted by $\overline{F}^3$ (L), $\overline{F}^3$ (L*) according as the pinch-points are distinct or coincident, (L) being the symbol indicating the presence of a double line.

§ 3. *The Generation, by Projection, of Certain Nodes.*

3*a*. *The projection of a surface from a point of a line thereon has in general a multiple point, the projection of the line.*

A criterion that a point shall be multiple (say, s-ple) on a surface in $[r - 1]$, is that every $[r - 2]$ through it shall meet the surface in a curve with an s-ple point there: so that to prove the above result it is sufficient to show that every $[r - 1]$ through the line of the original surface in $[r]$ meets the surface further in a curve intersecting the line in s points. Supposing for the moment that there are no multiple points of the surface on the line, these intersections are the points at which the tangent plane to the surface lies in the given $[r - 1]$, *i.e.*, s is simply the order of the line-cone formed by the ∞^1 tangent planes at the points of the line.

For an F^n this order is in general two [*v.* § 4*bii*].

Remarks.

3*aii*. The above argument is valid without modification when the line is multiple upon the surface. It will be observed that if the line is simple, the nodal cone obtained is necessarily rational. If there are points of the line, whose multiplicity upon the surface exceeds that of the line, s must be increased by the sum of the excesses [*cf.* § 3*b*].

3*aiii*. The above generation of a node may be regarded as the final stage in projecting a surface from a $[q - 1]$, lying in a $[q]$ which contains a curve of the surface. This is the general result.

3*aiv*. In projecting a surface from a simple point thereon, the adjacent points give rise to a line of the projected surface; thus the above is equivalent to: the projection of a surface from a tangent line has a (rational) node, the projection of the tangent plane; as may be seen directly [*cf.* del Pezzo, 'Rend. del Circ. Mat. di Palermo,' vol. 1, p. 265 (1887)].

3*a*v. The method can easily be applied to show that a node of a surface projects into a node, with no exceptions save when the point of projection lies at the node or, in certain circumstances, on a ray through it.

3*a*vi. The most important case for which $s = 1$, so that only a simple point is obtained, is that of a generator of a ruled surface.

3*b*. *The projection of a surface from a point on a ray, i.e., a line of the surface through a node, has in general a node, the projection of the ray, whose nodal cone is degenerate.*

For the points of the new surface, in the neighbourhood of the projection of the ray, are the projections of two algebraically distinct sets, viz.:—

(i) The points adjacent to the various points of the ray, defining the tangent planes.

(ii) The points adjacent to the node, lying on the nodal cone.

In the case of an F^n the point is simply biplanar, for the projections of (i) lie in a plane, the projection of the solid which, as shown in § 4*b*ii, contains the tangent planes at the points of the ray; whilst the projections of (ii) lie in a plane, the projection of the nodal solid.

Remarks.

3*b*ii. By combination with the previous result, such a node can be obtained by projection from (α) a tangent line at a point of a line on the surface to be projected, or (β) two points, on intersecting lines of the surface.

3*b*iii. Using 3*b*ii β, on an F^n, since the tangent planes at the points of the two lines lie in [5] [*v.* § 4*b*ii β], the biplanes lie in a solid.

3*b*iv. A binode may also be the projection of a bisecant (of the original surface) from a point thereon. On any surface in three dimensions, however, such a binode is necessarily a point of a double curve; for the bisecants of a surface in [4] which pass through a point, form a surface (viz., a cone) which cannot meet a solid in an isolated point [*cf.* Segre, 'Ency. d. math. Wiss.,' vol. III$_2$, 7, p. 907]. The contrary is implied by del Pezzo (*loc. cit.*, p. 249); the projections which he suggests produce conic-nodes [*v.* § 9*c*].

Such a generation of an isolated binode appears to be impossible on any F^n.

§ 4. *The General Properties of the* F^n.

4*a*. The F^n, *i.e.*, the non-ruled surfaces of order n in space of n dimensions, have been treated by del Pezzo,† who shows that such a surface projects, from n-3 general points upon it, birationally, and therefore into a cubic surface. That the tangent planes at two general points do not intersect, and thus that the lines of the cubic surface, which are the projections of the tangent planes at the points of projection O, are skew ; since a cubic surface has no sets of skew lines in number more than six, $n \leqslant 9$. That all the surfaces are projections of the F^9—which is found to be necessarily non-singular—save F^{8*}, for which the lines of the cubic surface arising from the tangent planes at the five points O are such that there are upon the cubic surface two lines meeting all the five and no line skew to all five. Since a curve of the cubic surface meeting l of the appropriate set of skew lines must be the projection of a curve through l of the points O, whose order exceeds by l that of the projected curve, it is possible to determine the properties of these surfaces. And that all these surfaces are representable in the plane ; all, save F^{8*} and those distinguished as $\overline{F}^n$ [*v.* § 2*e*], by the cubics through $9-n$ points, and F^{8*} by the quartics with two fixed nodes. Hereafter until § 7 the $\overline{F}^n$ will be excluded : the cubics which represent them are nodal or degenerate and not completely determined by their fixed points. It may be noted that ruled surfaces (including rational cones) of order n in space of n dimensions, require in general curves of higher order to represent them ; and that elliptic cones are not representable.

4*b*i. The plane representation affords a convenient method for investigating the curves upon an F^n, least satisfactory in the case of the lines on a surface with a complicated singularity. It is easy to show, from merely numerical considerations relating to the systems of curves in the plane, that, apart from points on their lines, the F^n are homogeneous in the sense that if a curve, on the surface, of a given type passes through one point, such a curve can be drawn to pass through any point.

It may be shown by similar reasoning that any F^4—but not a general $\overline{F}^4$, is the intersection of two quadrics in [4].

4*b*ii. In projecting the F^n the following theorem is useful :—

(α) *The tangent planes at the points of a line, ray or axis respectively of an* F^n *form line-cones of order* 2, 1, 0, *and therefore lie in* [4], [3], [2]. [Terms explained in § 2*d*.]

(β) The tangent planes at the points of a conic (in particular of two intersecting lines), a cubic (in particular of two skew lines and a transversal) or a quartic, of an F^n, lie respectively in [5], [6], [7].

---

† 'Rend. del. Circ. Mat. di Palermo,' vol. 1, p. 241 (1887).

Proof.—(α) After projection from $n-4$ general points of the surface, it evidently suffices to prove the statement for F^4, *i.e.*, as remarked in § 4*b*i, the intersection of two quadrics in [4].

The tangent primes to a quadric of [4], at the points of a line thereon, form a pencil (*i.e.*, contain a fixed plane); so that their intersections with an arbitrary plane are the lines of a pencil related to the points of the line on the quadric.

The tangent plane at a point of a line of an F^4 is the intersection of the tangent primes to the quadrics, at that point.

Thus the various such tangent planes meet an arbitrary plane in the points which are the intersection of corresponding members of two related pencils, *i.e.*, lie on a conic. If, however, the quadrics touch at a point of the line (the point of contact being, of course, a node on their intersection) so that there is a self-corresponding line, the conic reduces to a line; and if the quadrics touch twice, there are two self-corresponding lines, so that the pencils will have the same centre, and all the tangent planes meet the arbitrary plane in this point, *i.e.*, coincide. This is equivalent to the enunciation (α).

(β) This is reduced to (α) by projecting the curves into lines; a modification is required for the degenerate cases, as follows:—

Project an F^n from the intersection O of two of its lines: these project into nodes [*v.* § 3*a*], whose join is the projection of the tangent plane at O. Along this axis the tangent plane is fixed [see (α)], and is common to the two nodal solids, *i.e.*, the solids containing the nodal cones, so that the tangent planes to these cones lie in [4] (= [3 + 3 − 2]).

Hence the tangent planes at the points of the two intersecting lines of the original surface lie in a [5].

The tangent planes at the points of two lines and a transversal lie in two [5]s with a [4] common, *i.e.*, in a [5 + 5 − 4] = [6].

§ 5. *The Generation of the Special Nodes of* F^n.

5*a*. The various nodes of the F^n are generated as the projections of simpler singularities (and thus by combination, of simple points only) as follows [for notation *v.* § 2*b*, *c*, *d*]:—

(C_2) as the projection of (A) from a point on a line of the surface [§ 3*a*].

(B_3) as the projection of (C_2) from a point on a ray through C_2 [§§ 3*b* and 5*c*].

(B_4) as the projection of $(2C_2)$ from a point on the axis joining the nodes [§ 5*d*].

(B_4) or (B_4^*), according to circumstances, as the projection of (B_3) from a point on a ray through B_3 [§§ 5*e*, 5*f*].

(B_4) as the projection of (B_4^*) from a general point of the surface [§ 5*f*].

(B_5) as the projection of
- $(B_3 + C_2)$ from a point on the axis joining the nodes [§ 5*g*].
- (B_4) from a point on a ray through B_4 [§ 5*h*].
- (B_4^*) from a point on a ray through B_4^* [§ 5*i*].

(B_6) as the projection of
- $(B_4^* + C_2)$ from a point on the axis joining the nodes [§ 5*j*].
- (B_5) from a point on a ray through B_5 [§ 5*k*].

(U_6) as the projection of (B_4) from a point on the edge [§ 5*l*].

(U_7) as the projection of
- (B_5) from a point on the edge [§ 5*m*].
- (U_6) from a point on a ray through U_6 [§ 5*n*].

(U_8) as the projection of (U_7) from a point on the torsal ray through U_7 [§ 5*p*].

For each of these results a proof will now be given, general in character, but, for definiteness, stated as relating to a particular one of the surfaces.

5*b*. As has already been assumed, the order of a curve or other variety (and the dimension of a flat space) is reduced by projection from a point O of itself. In the case of a surface the points adjacent to O give rise to a line of the projected surface whose points correspond to the directions through O, so that curves through O project into curves meeting this line in points which will be distinct unless the curves on the original surface touch at O.

5*c*. $(B_3) = (C_2)$ — (point on ray) exemplified by $F^6\,(B_3) = F^7\,(C_2) - (b)$.

The ray b on which the point of projection O stands projects into a binode, whose biplanes β, β' are the projections respectively of the tangent-plane-solid along b, and of the nodal solid. The line d and the tangent plane at O project into rays b_1, b_2 both lying in the biplane β: these are the only lines of the surface. The conics of $F^7\,(C_2)$ pass through the node, and, therefore, since they lie on the surface, touch the nodal cone and consequently the nodal solid; so that their projections, the conics χ^2, pass through B_3 and touch the biplane β'. The quartics of $F^7\,(C_2)$ meeting b project into quartics K^4 which pass through B_3, touching the biplane β.

In particular, if the point of projection be taken at the point (b, d) the ray b projects into a binode, biplanes β, β' as before, but the line d into a conic-node, and the tangent plane at O into the axis joining the two nodes.

5*d*. $(B_4) = (2C_2)$ — (point on axis). Nothing is to be gained by considering a particular case.

The projection of an F^n from a point on an axis joining two conic-nodes has a B_4, the projection of the axis. For the points of the new surface adjacent to this point are the projections of the points adjacent to the nodes and lying on the nodal cones, and of the points of the surface adjacent to the axis, *i.e.*, of the fixed tangent plane along the axis [§ 4*b*ii α]: the former are points lying in two planes, the projections of the nodal solids, and the latter the intersection of these, the edge e, which therefore lies upon the surface.

$$5e. \left\{\begin{matrix} (B_4) \\ \text{or} \\ (B_4^*) \end{matrix}\right\} = B_3 - \text{(point on ray), exemplified by} \left\{\begin{matrix} F^6(B_3) - (b_1) = F^5(B_4). \\ \\ F^5(B_3) - (b') = F^4(B_4^*). \end{matrix}\right.$$

In the former case the line b_1 projects into a binode [*cf.* § 3*b*], whose biplanes β', β are the projections respectively of the binodal solid (in fact of the biplane β') of $F^6(B_3)$, and of the tangent-plane-solid along the ray b_1: their intersection is the projection of the biplane β of $F^6(B_3)$ which contains b_1, but this contains also the ray b_2, so that the edge e lies on the surface F^5.

The tangent planes at the points of the edge coincide, viz., in the projection of the tangent-plane-solid along b_2. β contains a ray b_1, the projection of the tangent plane at O.

5*f*. In the latter case, b' projects into a binode, biplanes β, β', the projections of β and of the tangent-plane-solid along b', the edge that of the biplane β', as in the former case; but, since β' contains no line of $F^5(B_3)$ other than b', the edge does not lie on the new surface, though the binode is obviously special, it will be denoted by B_4^*. The conics χ^2 which touch the biplane β' of $F^5(B_3)$ project into conics χ^2 touching the edge. On projection from any point O of $F^4(B_4^*)$, the conic χ^2 through O gives rise to a line of the projected surface F^3, touching, and therefore coinciding with, the edge; so that the binode of the F^3 is an ordinary B_4, which explains the perhaps surprising existence of a singularity on an F^4 with no complete analogue on any F^3.

Segre distinguishes B_4^* from B_4 as a coincident couple of conic-nodes [*cf.* § 6*g*iv].

5*g*. $(B_5) = (B_3 + C_2)$ — (point on axis) exemplified by $F^5(B_5) = F^6(B_3 + C_2) - (a)$.

$F^6(B_3 + C_2)$ regarded as a projection of $F^7(C_2)$ is the limiting case of $F^6(B_3)$

as O moves along b to its intersection with d, a being the limiting position of b_1 and b_2. Hence $F^6(B_3 + C_2) - (a)$ is the limiting case of $F^6(B_3) - (b_1)$ as b_2 tends to b_1, so that the projections of the tangent-plane-solids along these two rays coincide; since these projections are the fixed tangent plane along the edge and the biplane β (of $F^5(B_5)$), β becomes a torsal biplane.

5*h*. $(B_5) = (B_4)$ — (point on ray) exemplified by $F^4(B_5) = F^5(B_4) - (b)$.

b projects into a binode, biplanes β', β, the projections respectively of the biplane β' and of the tangent-plane-solid along b; the fixed tangent plane along the edge c projects into a fixed tangent plane along the edge, which, moreover, coincides with the biplane β', for both are projections of planes in the binodal solid, which contains the point of projection, *i.e.*, β' is a torsal biplane.

5*i*. $(B_5) = (B_4{}^*)$ — (point on ray) exemplified by $F^3(B_5) = F^4(B_4{}^*) - (b_1)$.

$F^4(B_4{}^*) - (b)$ is the limit of $F^4(B_4{}^*) - (P)$ as P approaches b_1; the ray b_2 of the biplane β projects into a ray b_2 of a biplane β, which in the limit coincides with the edge, the projection of the conic χ^2 through P, *i.e.*, β is a torsal biplane.

5*j*. $(B_6) = (B_4{}^* + C_2)$ — (point on axis) exemplified by $F^3(B_6) = F^4(B_4{}^* + C_2) - (a)$.

$F^4(B_4{}^* + C_2)$ is the limit of $F^4(B_4{}^*) - (b_1)$ when b_1, b_2 coincide, *i.e.*, the limit of $F^3(B_5)$ when the ray in the torsal biplane coincides with the edge, *i.e.*, this biplane is oscular.

5*k*. $(B_6) = (B_5)$ — (point on ray) exemplified by $F^3(B_6) = F^4(B_5) - (b)$.

This is proved by considering the limit of $F^4(B_5) - (P)$ as P approaches b.

5*l*. $(U_6) = (B_4)$ — (point on edge) exemplified by $F^4(U_6) = F^5(B_4) - (e)$.

The section of $F^5(B_4)$ by any [4] through the edge consists of the edge and a residual quartic through the node. This quartic touches the edge at the node; for any curve on the surface, through the binode, touches at least one of the biplanes, and the [4] which contains this quartic meets both biplanes in the edge only. Thus the surface obtained by projecting from a point on the edge is met by every [3] through the point which is the projection of the edge, in a curve having a cusp there, *i.e.*, the point is a unode. The uniplane is the projection of the binodal solid and contains rays b_1, b_2, the projections of b and of the fixed tangent plane along the edge; along each ray the tangent plane is fixed, *e.g.*, in the projection of the tangent-plane-solid along b.

5*m*. $(U_7) = (B_5)$ — (point on edge) exemplified by $F^4(U_7) = F^5(B_5) - (e)$.

$F^5(B_5)$ is the limit of $F^5(B_4)$ when b coincides with e; wherefore the F^4 obtained by the projection under consideration is the limit of $F^4(U_6)$ when b_1, b_2 coincide. The common limit of the fixed tangent planes along b_1, b_2 is their common plane, *i.e.*, the uniplane is torsal.

5n. $(U_7) = (U_6)$ — (point on ray); the sole example is $F^3(U_7) = F^4(U_6) - (b_1)$.

The residual section of $F^4(U_6)$ by any solid through b_1 is a cubic touching the section of the uniplane, *i.e.*, b_1; wherefore the section of the projected surface by a plane through the projection of b_1 has a cusp there, so that this point is a unode.

The section of $F^4(U_6)$ by a solid passing through the fixed tangent plane along b_1, and also through b_2, consists of b_1 counted twice, of b_2, and therefore further of a line, which can only be b_2 repeated. Thus this solid, which is in fact the nodal solid, contains both fixed tangent planes, and projects from a point of b_1 into a plane containing the projection of the tangent plane along b_1 simply, and the projection of b_2 doubly, *i.e.*, the uniplane of the projected surface is torsal along the projection of b_2.

5p. $(U_8) = (U_7)$ — (point on torsal ray); the sole example is $F^3(U_8) = F^4(U_7) - (b)$.

This is the limit of $F^4(U_6) - (b_1)$ when b_1, b_2 coincide, *i.e.*, the limit of $F^3(U_7)$ when the ordinary ray coincides with the torsal ray; so that the uniplane is oscular along this, the only line of the surface.

§ 6. *Description of the Nodal* F^n.

In this paragraph are described the various surfaces arising in the process proposed in § 1*a* (and tabulated in § 8). A set of diagrams is appended.

6a. The surfaces are treated with a view to their projection—ultimately into an F^3—and accordingly only the properties required for this purpose are enunciated. In particular, on account of the peculiar importance of the lines of the surfaces, those curves of each surface are described which may afterwards project into lines of another. A curve of order r will be denoted by K^r or χ^r, lower suffixes being employed to distinguish between different systems.

It is naturally not possible, in general, to determine the number of intersections of two curves of a surface from a knowledge merely of their orders; but the following rules, of fairly extensive applicability to the conics, cubics and quartics *actually considered* upon *any* F^n are convenient, the exceptions only (indicated in parenthesis) being noted in the sequel: the number of intersections is :—

Two conics, same system, 0; different systems, 1 (or 2).

Two cubics, same system, 1; different systems, 2.

Two quartics, same system, 2; different systems, 3.

Conic and cubic, 1 (or 2); conic and quartic, 1 (or 2); cubic and quartic, 2.

N.B.—(Fixed) intersections at nodes are excluded.

Every such rule is strictly valid provided n is great enough.

Each number in the arrays appearing in the descriptions of F^8 and F^7 is the number of intersections of the curves named at the head of the corresponding row and column.

To save space, syntactic completeness is neglected.

In the diagrams a conic-node is represented by a conventionalised sketch of the nodal cone. Biplanes and uniplanes are shaded; when they are torsal this is indicated by a fine line adjacent to the edge or ray. In the diagrams of F^5 (A), F^4 (C_2) and F^4 ($2C_2$) the actual intersections of the lines of the surface are marked.

F^9.

6*b*. F^9 (A).—The F^9 is represented by all the cubics of the plane, from which it is easily seen that the order of every curve upon it is a multiple of 3, and that those of the same order form a single system. The only systems to be considered (*i.e.*, order $\leqq 7$) are the cubics and sextics, ∞^2 and ∞^5 respectively, so that for these—and in fact for no others—the freedom is one less than the order (and because they and the systems derived from them by the projections have no multiple points, or base-points save at the multiple points of the surfaces, this is the case for these other systems also). The sextics intersect in four points and meet the cubics in two.

The surface, being homogeneous, can only be projected from a general point.

F^8.

6ci. F^8 (A).—The required curves upon the surface and their intersections are as follows:—

		d	K^5	K^6
One line	d	:	1	0
One system of conics	K^2	1	1	2
One „ cubics	K^3	0	2	2
One „ quintics ...	K^5	1	3	4
One „ sextics	K^6	0	4	4

6cii. F^8 (A^*).—It is easily found from the plane representation (by the quartics with two fixed nodes) that the required curves upon the surface are:—

		K_1^6	K_2^6
Two systems of conics ...	K_1^2	1	2
	K_2^2	2	1
One system of quartics ...	K^4	3	3
Two systems of sextics ...	K_1^6	4	5
	K_2^6	5	4

As in the case of the required systems of curves upon F^9, the freedom is one less than the order. The surface is homogeneous.

F^7.

6*d*i. F^7 (A).

		d'	d_1	d_2	K_1^5	K_2^5
Three lines	d'	·	1	1	1	1
	d_1	1	·	0	0	1
	d_2	1	0	·	1	0
Two systems of conics ...	K_1^2	0	0	1	1	2
	K_2^2	0	1	0	2	1
One system of cubics	K^3	1	0	0	2	2
One ,, quartics	K^4	0	1	1	3	3
Two systems of quintics ...	K_1^5	1	0	1	3	4
	K_2^5	1	1	0	4	3

The projection from the line d' is the Veronese surface, from d_1 or d_2 a ruled surface.

6*d*ii. F^7 (C_2).

		C	a	b	χ^5
The node	C	·	1	0	1
Two lines	b	1	·	1	0
	d	0	1	·	1
One system of conics	χ^2	1	0	0	1
One ,, cubics	K^3	0	0	1	2
One ,, quartics ...	K^4	0	1	0	3
One ,, quintics ...	χ^5	1	0	1	3

F^6.

6*e*i. F^6 (A).—The lines upon the surface form a hexagon (they are the projection of d_1, d', d_2, the K_2^2 through O, the tangent plane at O, the K_1^2 through O).

Three systems of conics : K_1^2, K_2^2, K_3^2 ; each meets two opposite sides of the hexagon.

Two ,, cubics : K_λ^3, K_μ^3 ; each meets three alternate sides.

Three ,, quartics : K_1^4, K_2^4, K_3^4 ; each meets the four sides left on removing two opposite.

K_i^2, K_i^4 meet in two points.

6*e*ii. F^6 (C_2).

A conic-node C.

Two rays b_1, b_2 met respectively by the lines d_1, d_2.

Two systems of conics : K^2, χ^2. K^2 meets d_1, d_2 ; χ^2 passes through C.

Two ,, cubics : K_1^3, K_2^3. K_1^3 meets b_1, d_2 ; K_2^3 meets b_2, d_1.

Two ,, quartics : K^4, χ^4. K^4 meets b_1, b_2 ; χ^4 meets d_1, d_2 and passes through C.

K^2, K^4 ; and χ^2, χ^4 meet in two points.

6*e*iii. F^6 (C_2^*).

A conic-node C^*.

Three rays b_1, b_2, b_3.

Three systems of conics : K_1^2, K_2^2, K_3^2. K_1^2 meets b_1, etc.

One system of cubics : χ^3, passes through C^*.

Three systems of quartics : K_1^4, K_2^4, K_3^4. K_1^4 meets b_2, b_3, etc.

K_i^2, K_i^4 meet in two points.

The node C_2^* of this surface (and of the succeeding) is the only node of an F^n through which there does not pass a system of conics, and consequently the only point from which an F^n can be projected into a non-ruled surface of order $n-2$ in $n-1$ dimensions, viz., the Veronese surface, of order 4 in 5 dimensions [*v.* § 7*a.*].

From a general point of the surface $F^6(C_2)$ and $F^6(C_2^*)$ project into the same $F^5(C_2)$.

6*e*iv. $F^6(2C_2)$.

Two conic-nodes C, C^*, similar to those of the preceding two surfaces, joined by an axis a.

Through C^* passes a ray b.
Two systems of conics : K^2, χ^2. K^2 meets b ; χ^2 passes through C.
One system of cubics : χ^3 passes through C^*.
Two systems of quartics : K^4, χ^4. K^4 meets a, χ^4 meets b and passes through C.
K^2, K^4 meet in two points.

6*e*v. $F^6(B_3)$.

A binode B_3, biplanes β, β'.
Two rays b_1, b_2 in the biplane β.
One system of conics χ^2, passing through B_3, touching β'.
Two systems of cubics K_1^3, K_2^3, meeting b_1, b_2 respectively.
One system of quartics K^4, passing through B_3, touching β.

6*e*vi. $F^6(B_3 + C_2)$.

A binode B_3, biplanes β, β'.
A conic-node C.
An axis a, joining B_3, C, lying in the biplane β.
One system of conics, χ^2, passing through B_3, touching β'.
One ,, cubics K^3, ,, C.
One ,, quartics K^4, ,, B_3, touching β.

F^5.

6*f*i. $F^5(A)$.—The 10 lines from a well-known configuration, intersecting as do the lines of a cubic surface which meet neither of two skew lines thereon.

Five systems of conics K_i^2, and five corresponding systems of cubics K_i^3 $(i = 1, 2, 3, 4, 5)$.
K_i^2 meets the four lines *ij*, *ik*, *il*, *im*, and K_i^3 the remaining six.
K_i^2, K_i^3 meet in two points.

6*f*ii. $F^5(C_2)$.

A conic-node C.
Three rays b_1, b_2, b_3, met respectively by the lines d_1, d_2, d_3.
All of the latter three meet a line d'.
Four systems of conics K_1^2, K_2^2, K_3^2, χ^2 ; four systems of cubics K_1^3, K_2^3, K_3^3, χ^3.
K_1^2 meets b_1, d_2, d_3 ; χ^2 meets d' and passes through C.
K_1^3 meets d_1, b_2, b_3, d' ; χ^3 meets d_1, d_2, d_3 and passes through C.
K_i^2, K_i^3 meet in two points.

6*f*iii. $F^5(2C_2)$.

Two conic-nodes $C_{(1)}$, $C_{(2)}$.

The lines form a pentagon, consisting of the axis a joining the nodes $C_{(1)}$, $C_{(2)}$, the rays b_1, b_2 passing respectively through $C_{(1)}$, $C_{(2)}$ and the intersecting lines d_1, d_2 which meet respectively b_1, b_2.

Three systems of conics K_1^2, K_2^2, χ^2 ; three systems of cubics K_1^3, K_2^3, χ^3.
K_1^2 passes through $C_{(1)}$ and meets d_2 ; χ^2 meets b_1, b_2.
K_1^3 passes through $C_{(1)}$ and meets d_1, b_2 ; χ^3 meets a, d_1, d_2.
χ^2, χ^3 meet in two points.

6*f*iv. $F^5(B_3)$.

A binode B_3, biplanes β, β'.

In the biplane β two rays b_1, b_2; in the biplane β' one ray b', met by a line d.

Three systems of conics K_1^2, K_2^2, χ^2; three systems of cubics K_1^3, K_2^3, χ^3.
K_1^2 meets b_1, d; χ^2 passes through B_3 touching β'.
K_1^3 meets b_1, b'; χ^3 passes through B_3, touching β, and meets d.
K_1^2, K_2^3; and K_2^2, K_1^3 meet in two points.

6*f*v. $F^5(B_3 + C_2)$.

A binode B_3, biplanes β, β'. A conic-node C.

In the biplane β the axis a joining B_3, C; in the biplane β' one ray b met by a line d.

Two systems of conics K^2, χ^2; two systems of cubics K^3, χ^3.
K^2 passes through C, and meets d.
χ^2 passes through B_3, touching β'.
K^3 passes through C, and meets b.
χ^3 passes through B_3, touching β, and meets d.

6*f*vi. $F^5(B_4)$.

A binode B_4, biplanes β, β'.
The edge e; a ray b, lying in β.
Two systems of conics K^2, χ^2; two systems of cubics K^3, χ^3.
K^2 meets b. K^3 meets e; χ^2, χ^3 pass through B_4, touching β', β respectively.
K^2, χ^3 meet in two points.

6*f*vii. $F^5(B_5)$.

A binode B_5, torsal biplane β, ordinary biplane β'.
The edge e.
One system of conics χ^2; one system of cubics χ^3.
χ^2 passes through B_5, touching the ordinary biplane β'.
χ^3 passes through B_5, touching the torsal biplane β.

F^4.

Each of these surfaces is the intersection of a pencil of quadrics in [4], [*v.* § 4*b*i]. They are treated from this standpoint by Segre ['Math. Ann.,' vol. 24, p. 313 (1888)]. For each case the invariant factors are given.

6*g*i. F^4(A). [11111.]

The 16 lines intersect as do those of a cubic surface when any line (say 6) and the 10 which meet it are abstracted.

There are five pairs of systems of conics, two conics of a pair meeting complementary sets of eight lines (*e.g.*, 1, 2, 3, 4, 15, 25, 35, 45 and 23, 31, 12, 14, 24, 34, 5, 6′) and intersecting in two points.

6*g*ii. $F^4(C_2)$. [2111.]

Twelve lines: four rays b_1, b_2, b_3, b_4, each met by two mere lines, *e.g.*, d_1, d_1'; these form a pair of conjugate lines in a "double-four" (*i.e.*, b_i, d_j—or b_i, d_j'—intersect if $i = j$; d_i, d_j skew; d_i, d_j' intersect if $i \neq j$).

Eight systems of conics.

Six systems K_{ij}^2 each meeting two rays b_i, b_j and the four lines skew to both, *i.e.*, d_k, d_l, d_k', d_l'.

Two systems K^2, K'^2 each meeting four skew lines and passing through the node.

K_{ij}^2, K_{kl}^2 meet in two points.

*6g*iii. $F^4(2C_2)$. [221.]

Two conic-nodes C, C'.

Nine lines: the axis a; two rays through each node—b_1, b_2 through C, b_3', b_4' through C'; four lines, each meeting one ray through C and one ray through C' (viz., $13'$, $14'$, $23'$, $24'$), of which two meeting no ray in common (*i.e.*, $13'$, $24'$ and $14'$, $23'$) intersect.

Six systems of conics.

Two systems through each node (K_3^2, K_4^2 through C, $K_1'^2$, $K_2'^2$ through C'), meeting one ray through the other node and the two mere lines skew thereto (*e.g.*, $K_2'^2$ through C' meets b_2, $13'$, $14'$).

One system K_ρ^2 meeting all four rays.

One system K_λ^2 meeting all four mere lines and the axis.

K_ρ^2, K_λ^2 meet in two points.

*6g*iv. $F^4(2C_2^*)$. [(11)111.]

Through the node C four rays b_1, b_2, b_3, b_4 met respectively by b_1', b_2', b_3', b_4' through the node C'. The join of C and C' does *not* lie on the surface.

Seven systems of conics.

Six systems K_{ij}^2 each meeting two rays through one node and the two through the other node skew to both, *i.e.*, b_i, b_j, b_k', b_l'.

One system χ^2 passing through both nodes.

K_{ij}^2, K_{kl}^2 meet in two points.

From a general point of the surface, $F^4(2C_2)$ and $F^4(2C_2^*)$ project into the same $F^3(2C_2)$.

Segre [*loc. cit.*, p. 380] calls two nodes whose join does not lie on the surface a couple of nodes. $F^4(3C_2)$, $F^4(B_3+2C_2^*)$, $F^4(B_4^*+2C_2^*)$ each contain a couple. The four nodes of $F^4(4C_2)$ consist of two couples. B_4^* may be regarded as a coincident couple.

*6g*v. $F^4(3C_2)$. [(11)21.]

Six lines.

The node $C_{(0)}$ is joined to the nodes $C_{(1)}$, $C_{(2)}$ by the axes a_1, a_2.

Through each of the nodes of the couple pass two rays, b_1, b_1' through $C_{(1)}$; b_2, b_2' through $C_{(2)}$; b_1 meets b_2', b_2 meets b_1'.

Five systems of conics.

Two systems K_1^2, K_2^2, each meeting an axis and the rays through the node on the other axis, *e.g.*, K_1^2 meets a_1, b_2, b_2'.

Two systems K^2, K'^2 passing through $C_{(0)}$ and meeting 1, 2; 1′, 2′ respectively.

One system χ^2 passing through $C_{(1)}$, $C_{(2)}$.

K_1^2, K_2^2 meet in two points.

*6g*vi. $F^4(4C_2)$. [(11)(11)1.]

The four nodes, $C_{(1)}$, $C_{(2)}$, $C_{(3)}$, $C_{(4)}$ are the vertices of a quadrilateral (a_{12}, etc.), lying on the surface.

Four systems of conics.

Two systems $K^2_{(12,\,34)}$, $K^2_{(14,\,23)}$ each meeting two opposite axes.

Two systems K_{13}^2, K_{24}^2 each passing through two opposite nodes.

$K^2_{(12,\,34)}$, $K^2_{(14,\,23)}$ meet in two points.

*6g*vii. $F^4(B_3)$. [311.]

Eight lines: two rays in each biplane (b_1, b_2 in β; b_4', b_5' in β') each met by one line (d_1, d_2, d_4', d_5' respectively); lines meeting rays in different biplanes intersect.

Six systems of conics.

Four systems (K_{14}^2, K_{15}^2, K_{24}^2, K_{25}^2), each meeting two rays, in different biplanes (indicated by suffix) and the two lines skew to both.

Two systems χ^2, χ'^2 passing through B_3, touching the biplanes β, β' and meeting the lines d_4', d_5'; d_1, d_2 respectively.

K_{14}^2, K_{25}^2; and K_{24}^2, K_{15}^2 meet in two points.

6*g*viii. $F^4(B_3 + C_2)$. [32.]

Six lines: in the biplane β' the axis a; in the biplane β two binodal rays b_1, b_2 meeting respectively the lines d_1, d_2; a conic-nodal ray b' meets d_1, d_2.

Four systems of conics χ^2, χ'^2, K_1^2, K_2^2.

χ^2 passes through B_3, touching β, and meets b'.

χ'^2 passes through B_3, touching β', and meets d_1, d_2.

K_1^2 passes through C, and meets b_1, d_2.

6*g*ix. $F^4(B_3 + 2C_2^*)$. [3(11).]

B_3 is joined to $C_{(1)}$, $C_{(2)}$ by the axes a_1, a_2 lying in the biplanes β_1, β_2 respectively; through the conic-nodes pass respectively the intersecting rays b_1, b_2.

Three systems of conics K_1^2, K_2^2, χ^2.

K_1^2 passes through B_3, touching β_1, and meets b_2.

χ^2 passes through both conic-nodes.

6*g*x. $F^4(B_4)$. [41.]

Five lines: the edge e; in each biplane (β_1, β_2) a ray (b_1, b_2) meeting the intersecting lines d_1, d_2 respectively.

Four systems of conics χ_1^2, χ_2^2, K_ρ^4, K_λ^2.

χ_1^2 passes through B_4, touching β_1, and meets b_2.

K_ρ^2 meets the two rays.

K_λ^2 meets the edge and the two mere lines.

K_ρ^2, K_λ^2 meet in two points.

6*g*xi. $F^4(B_4^*)$. [(21)11.]

A binode B_4^* [*v.* § 5*f*], or coincident couple of conic-nodes; biplanes β, β'.

In each biplane two rays; b_1, b_2 in β; b_5', b_6' in β'.

Five systems of conics.

One system χ^2 passes through B_4^*, touching the edge there.

Four systems K_{15}^2, K_{16}^2, K_{25}^2, K_{26}^2 meet one ray in each biplane (as indicated by suffix).

6*g*xii. $F^4(B_4^* + C_2)$. [(21)2.]

Three lines; in the biplane β' the axis a; in β two binodal rays b_1, b_2.

Three systems of conics.

χ^2 passes through B_4^*, touching the edge.

K_1^2, K_2^2 pass through C, and meet b_1, b_2 respectively.

6*g*xiii. $F^4(B_4^* + 2C_2^*)$. [(21)(11).]

B_4^* is joined to the conic-nodes $C_{(1)}$, $C_{(2)}$, by axes a_1, a_2 in different biplanes.

Two systems of conics.

χ^2 passes through B_4^*, touching the edge.

K^2 passes through both conic-nodes.

6*g*xiv. $F^4(B_5)$. [5.]

A binode B_5, torsal biplane β', ordinary biplane β.

Three lines: the edge e; a ray b, lying in β, and met by a line d.

Two systems of conics.

K^2 passes through B_5, touching β.

χ^2 passes through B_5, touching β', and meets d.

6*g*xv. $F^4(U_6)$. [(31)1.]

Two rays b_1, b_2 in the uniplane.

Three systems of conics.

K_1^2, K_2^2 meet b_1, b_2 respectively and intersect in two points.

χ^2 passes through U_6.

6*g*xvi. $F^4(U_7)$. [(41).]

The uniplane is torsal along the ray b.

A system of conics χ^2 passes through U_7.

F^3.

These surfaces are as given by Cayley [*v.* § 2*a*], though the notation for the lines is modified—in many cases only to the extent of suffixing his purely numerical symbols to *a*, *b*, *e* and *d*.

6*h*i. $F^3(A)$.—The familiar 27 lines.

6*h*ii. $F^3(C_2)$.—The configuration of 21 lines is well known in connection with Cremona's proof of Pascal's theorem, being that of $F^3(A)$ when $1 \equiv 1'$, $2 \equiv 2'$, …, $6 \equiv 6'$ (these six lines are rays).

6*h*iii. $F^3(2C_2)$.—Sixteen lines.

The axis a_0.

Through either node four rays, each meeting a corresponding ray through the other node (b_1, b_2, b_3, b_4; b_1', b_2', b_3', b_4').

Any two rays through one node and the two skew rays through the other are met by a line, which meets the similar line associated with the remaining four rays (*e.g.*, b_1, b_2, b_3' b_4' are met by (12 . 3′4′) which meets (34 . 1′2′)).

A mere line d_5 meets a_0 and the previous six mere lines.

6*h*iv. $F^3(3C_2)$.—Twelve lines.

Three axes a_1, a_2, a_3; a_1 joins $C_{(2)}$, $C_{(3)}$, etc.

Through each node two rays (b_1, b_1'; b_2, b_2'; b_3, b_3'): b_1 meets b_2', b_3', etc.

Three coplanar lines d_1, d_2, d_3; d_1 meets a_1, b_1, b_1', etc.

6*h*v. $F^3(4C_2)$.—Nine lines.

Six axes a_{ij}, joining the nodes in pairs ($i, j = 1, \ldots, 4$).

Three coplanar lines (ij . kl) meeting pairs of skew axes (a_{ij} and a_{kl}).

6*h*vi. $F^3(B_3)$.—Fifteen lines.

Six rays: b_1, b_2, b_3 in one biplane; b_4', b_5', b_6' in the other.

Nine mere lines, each meeting one ray in each biplane (14, 15, …, 36): ij meets kl if i, j, k, l all different.

6*h*vii. $F^3(B_3 + C_2)$.—Eleven lines.

The axis a_0.

In the axial biplane a ray b_1; in the ordinary biplane three rays b_2, b_3, b_4, met respectively by the three conic-nodal rays b_2', b_3', b_4'.

In each of the three planes determined by the axis and a binodal ray, one mere line, which meets also the two conic-nodal rays skew to the said binodal ray (12 . 3′4′, 13 . 2′4′, 14 . 2′3′).

6*h*viii. $F^3(B_3 + 2C_2)$.—Eight lines.

Two axes a_1, a_2, in the biplanes β_1, β_2 respectively, joining the binode to the conic-nodes $C_{(1)}$, $C_{(2)}$.

An axis a_0 joining the conic-nodes.

In each biplane one ray (b_1, b_2).

Through each conic-node one ray (b_1', b_2'); b_1 meets b_2', b_2 meets b_1'.

A line d meets a_0, b_1, b_2.

6*h*ix. $F^3 (2B_3)$.—Seven lines.

The axis a_0, lying in a biplane common to the two binodes.

In the non-axial biplane of each node three rays, each meeting a corresponding ray through the other node: b_1, b_2, b_3; b_1', b_2', b_3'.

6*h*x. $F^3 (2B_3 + C_2)$.—Five lines.

An axis a_0 joining the binodes, and lying in their common biplane.

Two axes a_1, a_2, joining the conic-node to the binodes $B_{(1)}$, $B_{(2)}$ respectively.

In each biplane containing the conic-node a ray of the corresponding binode: these intersect.

6*h*xi. $F^3 (3B_3)$.—Three axes, each joining two nodes, and lying in their common biplane.

6*h*xii. $F^3 (B_4)$.—Ten lines.

The edge e_3.

In each biplane two rays: b_1, b_2 in β; b_5', b_6' in β'.

Any two rays in different biplanes are met by a line (15′, 16′, 25′, 26′); 15′ meets 26′, 16′ meets 25′.

A line d_4 meets the edge and these four mere lines.

6*h*xiii. $F^3 (B_4 + C_2)$.—Seven lines.

The axis a_0; the edge e_3.

Two rays b_1, b_2 in the non-axial biplane, meeting conic-nodal rays b_1', b_2' respectively.

A line 12 meets e_3, b_1', b_2'.

6*h*xiv. $F^3 (B_4 + 2C_2)$.—Five lines.

Two axes, a_1, a_2, joining B_4 to the other two conic-nodes.

An axis a_0 joining the two conic-nodes.

The edge e_3.

A line 4 meets e_3 and a_0.

6*h*xv. $F^3 (B_5)$.—Six lines.

The edge e_0.

In the torsal biplane the ray b_1; in the ordinary biplane the rays b_2', b_3'.

Two lines 12′, 13′, each meeting b_1 and one ray of the ordinary biplane.

6*h*xvi. $F^3 (B_5 + C_2)$.—Four lines.

The axis; the edge; a ray in the torsal biplane; a conic-nodal ray: the two rays intersect.

6*h*xvii. $F^3 (B_6)$.—Three lines.

The edge, and two rays in the ordinary biplane.

6*h*xviii. $F^3 (B_6 + C_2)$.—Two lines.

The edge and the axis.

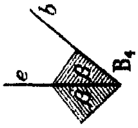
6*h*xix. $F^3 (U_6)$.—Six lines.

Three rays in the uniplane each met by one of three coplanar lines.

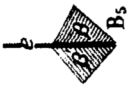
6*h*xx. $F^3 (U_7)$.—Three lines.

The torsal ray; the ordinary ray, met by a mere line.

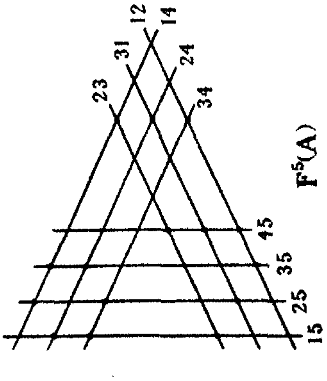
6*h*xxi. $F^3 (U_8)$.—One oscular ray.



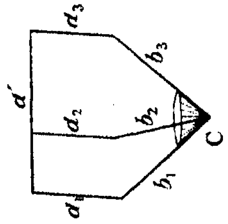
$F^5(B_4)$



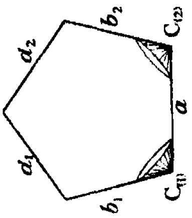
$F^5(B_5)$



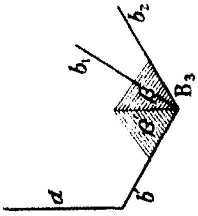
$F^5(A)$



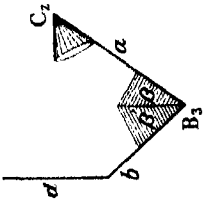
$F^5(C_2)$



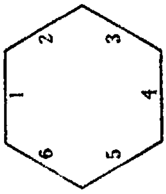
$F^5(2C_2)$



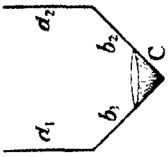
$F^5(B_3)$



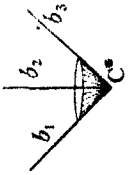
$F^5(B_3+C_2)$



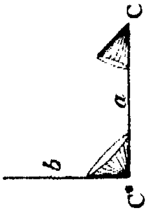
$F^6(A)$



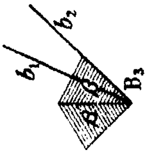
$F^6(C_2)$



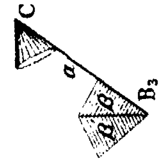
$F^6(C_2^*)$



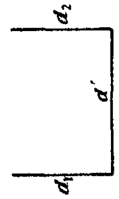
$F^6(2C_2)$



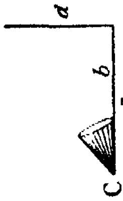
$F^6(B_3)$



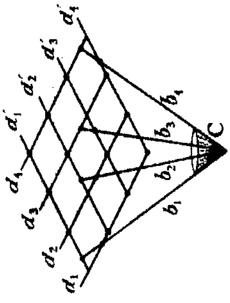
$F^6(B_3+C_2)$



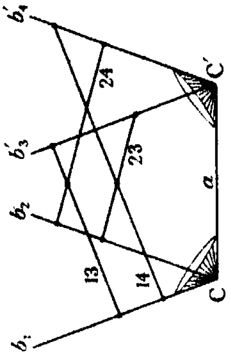
$F^7(A)$



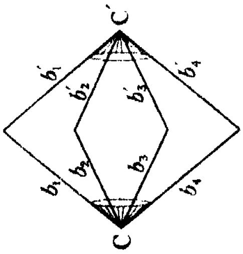
$F^7(C_2)$



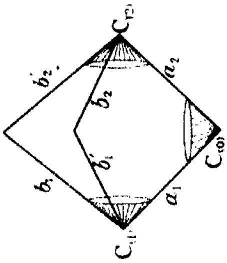
$F^4(C_2)$



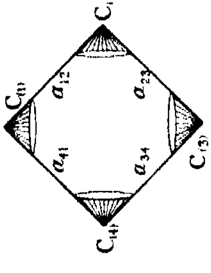
$F^4(2C_2)$



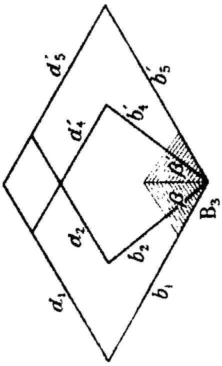
$F^4(2C_2^*)$



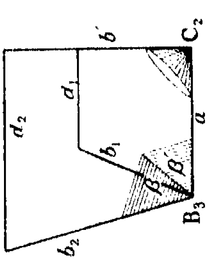
$F^4(3C_2)$



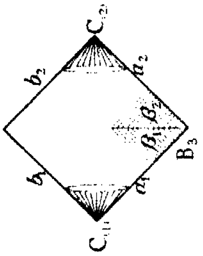
$F^4(4C_2)$



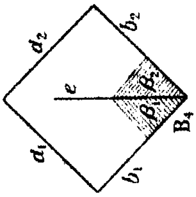
$F^4(B_3)$



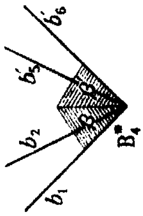
$F^4(B_3+C_2)$



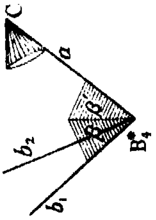
$F^4(B_3+2C_2^*)$



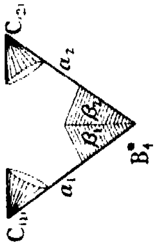
$F^4(B_4)$



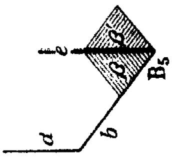
$F^4(B_4^*)$



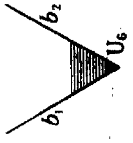
$F^4(B_4^*+C_2)$



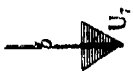
$F^4(B_4^*+2C_2^*)$



$F^4(B_5)$



$F^4(U_6)$



$F^4(U_7)$

§ 7. *The Projections of the Veronese Surface.*

7*a*. In the successive generation of F^n's a certain restriction has hitherto been observed, viz., that the points of projection shall be simple points of the surfaces —tantamount to assuming that they are represented in the plane by the complete system of cubics through $9-n$ points [§§ 1*a*, 4*a*]. This restriction being removed, it is necessary to consider what non-ruled surfaces F^n can be obtained by projecting an F^n from a node, and therefore—since this reduces the dimensions of the containing space by one, but the order by two—subsequently from an external point.

Because the projected surface is not to be ruled, no system of conics passes through the node, which is therefore the C^* of $F^6(C_2^*)$ or $F^6(2C_2)$. The projections from this node, and also those of $F^7(A)$ and $F^7(C_2)$ from d' and d respectively, are easily seen to be equivalent to the projection of F^9 from the space containing a twisted cubic thereon. The projected surface contains ∞^2 conics (the projections of the remaining cubics of F^9) and is the Veronese surface (of order 4 in [5]), of which, therefore, all the formerly excluded F^n are projections, as stated in § 1*a*.

The Veronese surface will be denoted by $\bar{\bar{F}}_V$ and its projections of order n in $[n]$, by $\bar{\bar{F}}^n$ (even if ruled).

7*b*. $\bar{\bar{F}}^4$.—The projection of the Veronese surface $\bar{\bar{F}}_V$ from a general external point O has no double point, for no chord of $\bar{\bar{F}}_V$ passes through O. But if O lies on the fourfold locus of chords, denoted by M^3, the projection has a double line, viz., the projection of the conic K of $\bar{\bar{F}}_V$ whose plane contains O. These two surfaces are denoted (in accordance with the notation of §§ 2*d*, 2*e*) respectively by $\bar{\bar{F}}^4(A)$, $\bar{\bar{F}}^4(L)$. The latter has two pinch-points, the projections of the points of contact of the tangents through O to the conic K.

Like the various F^4 considered in § 6*g*, $\bar{\bar{F}}^4(L)$ is the intersection of a suitable pencil of quadrics. $\bar{\bar{F}}^4(A)$ is not.

The only other projection of $\bar{\bar{F}}_V$ from a point (viz., upon it) is a ruled cubic.

7*c*i. The $\bar{\bar{F}}^3$, projections of $\bar{\bar{F}}^4$ from a point O thereon, are ruled, for ∞^1 conics pass through O; these are conveniently treated as particular projections of $\bar{\bar{F}}_V$ from a line.

The projection of $\bar{\bar{F}}_V$ from a line not meeting it is a Steiner surface [Bertini, 'Intro. alla Geom. proiett. degli ipersp.,' p. 422 (c. 16, § 18)], the double lines being the projections of the conics K_1, K_2, K_3 whose planes contain the points O_1, O_2, O_3 in which the line meets M^3: the pinch-points arise as on $\bar{\bar{F}}^4(L)$. There are 3, 2 or 1 distinct double lines according as the line is a proper trisecant, a tangent, or an inflexional tangent, of M^3. If the line lies upon M^3, there is an infinity of double lines, so that the projected surface is a repeated quadric.

7*c*ii. The line-vertex of projection may meet $\overline{F}_V$, which is double on M^3. The ∞^1 conics of $\overline{F}_V$ through this intersection O_{12} project into lines, so that the projected surface is a ruled cubic. The conic K_3 associated with O_3, the remaining intersection with M^3, projects into a double directrix having two necessarily distinct pinch-points [*cf.* $\overline{F}^4$ (L)]; the tangent plane at O_{12} into a simple directrix. Thus the projected surface is $\overline{F}^3$ (L).

7*c*iii. The line-vertex of projection may lie on the nodal cone of M^3 at a point O of $\overline{F}_V$; this cone is formed by the ∞^1 solids, defined by the tangent plane to $\overline{F}_V$ and the ∞^1 planes of the various conics of $\overline{F}_V$ through O. Hence the line of projection lies in a solid containing the points of $\overline{F}_V$ adjacent to O, and also the points of a conic of $\overline{F}_V$, which then together project into a double line whose only pinch-point is the projection of the tangent at O to the conic. Thus the projected surface is $\overline{F}^3$ (L*).

It will be observed that $\overline{F}^3$ (L) is a projection of $\overline{F}^4$ (A) or of $\overline{F}^4$ (L); $\overline{F}^3$ (L*) of $\overline{F}^4$ (A) only.

$\overline{F}_V$ projects from a chord into a quadric and from a tangent into a quadric cone.

§ 8. *Summary of Results.*

8*a*. *All cubic surfaces, including scrolls, have now been generated by the projection of non-singular surfaces, viz., the non-ruled surfaces of order n in n dimensions; indeed by the projection of the same surface* F^9 (*or* F^8).

Since the rational cones are representable in the plane, this is possible for them also [*cf.* § 9*a*]: in fact, projecting F^6 (B_3) from any line through B_3 in the biplane β', the conics through B_3, which touch β', become lines through a point, the projection of β', so that the projected surface is a rational cubic cone in [4]; the projection of this from an external point is a rational cubic cone, nodal in general, cuspidal if the point of projection lie on the locus of the tangent planes.

An elliptic cone is obtained when the ruled surface formed by joining the corresponding points of two elliptic cubic curves in birational correspondence, which lie in skew planes in [5], is projected from a line in either of these planes.

8*b*. All cubic surfaces except $\overline{F}^3$ (L*) and the cones have been obtained by projecting non-ruled, but in general nodal, surfaces, the complete intersection of a pencil of quadrics in [4], [*v.* table and § 7]. Further, there are two ruled quartic surfaces in [4], each the intersection of a pencil of quadrics—invariant factors [(22) 1] and [(32)]—[Segre, *loc. cit.*], and projecting from a point thereon into $\overline{F}^3$ (L) and $\overline{F}^3$ (L*) respectively. Finally by joining an elliptic, nodal or cuspidal quartic curve lying in a solid, to a point outside this solid, there is obtained a cone, the intersection of a pencil of quadric cones, which projects from a point thereon into the corresponding type of cubic cone.

Thus every cubic surface, cones included, is the projection of the complete intersection of a suitable pencil of quadrics in four dimensions.

8*c*. The results for the F^n not stated in § 1*a* are sufficiently summarised in the table.

On the left-hand side of a column are given, for each surface, the various essentially different points thereon from which it may be projected (as is easily verified from the diagrams of § 6; for the symbolism see § 1*a*). On the right-hand side are the surfaces generated by these projections; they appear on the left-hand side of the succeeding column. The figure in [ ] indicates the flat space of least dimensions which contains a non-singular surface of which the surface under consideration is a projection.

F^9 projected into F^8.

$F^9(A)$	$-(P)$	$= F^8(A)$
[9]		

F^8 projected into F^7.

$F^8(A)$	$-(P)$	$= F^7(A)$
[8]	$-(d)$	$= F^7(C_2)$
$F^8(A^*)$	$-(P)$	$= F^7(A)$
[8]		

F^7 projected into F^6.

$F^7(A)$	$-(P)$	$= F^6(A)$
	$-(d_1)$	$= F^6(C_2)$
[7]	$-(d')$	$= F^6(C_2^*)$
	$-(d_1 d')$	$= F^6(2C_2)$
$F^7(C_2)$	$-(P)$	$= F^6(C_2)$
	$-(d)$	$= F^6(2C_2)$
[8]	$-(b)$	$= F^6(B_3)$
	$-(bd)$	$= F^6(B_3+C_2)$

F^6 projected into F^5.

$F^6(A)$	$-(P)$	$= F^5(A)$
[6]	$-(1)$	$= F^5(C_2)$
	$-(1, 6)$	$= F^5(2C_2)$
$F^6(C_2)$	$-(P)$	$= F^5(C_2)$
	$-(d_1)$	$= F^5(2C_2)$
[7]	$-(b_1)$	$= F^5(B_3)$
	$-(b_1 d_1)$	$= F^5(B_3+C_2)$
$F^6(C_2^*)$	$-(P)$	$= F^5(C_2)$
[7]	$-(b_1)$	$= F^5(B_3)$
$F^6(2C_2)$	$-(P)$	$= F^5(2C_2)$
[7]	$-(b)$	$= F^5(B_3+C_2)$
	$-(a)$	$= F^5(B_4)$
$F^6(B_3)$	$-(P)$	$= F^5(B_3)$
[8]	$-(b_1)$	$= F^5(B_4)$
$F^6(B_3+C_2)$	$-(P)$	$= F^5(B_3+C_2)$
[8]	$-(a)$	$= F^5(B_5)$

F^5 projected into F^4.

$F^5(A)$	$-(P)$	$= F^4(A)$
[5]	$-(15)$	$= F^4(C_2)$
	$-(15, 34)$	$= F^4(2C_2)$
$F^5(C_2)$	$-(P)$	$= F^4(C_2)$
	$-(d')$	$= F^4(2C_2^*)$
[6]	$-(d_1)$	$= F^4(2C_2)$
	$-(b_1)$	$= F^4(B_3)$
	$-(d_1 d')$	$= F^4(3C_2)$
	$-(b_1 d_1)$	$= F^4(B_3+C_2)$
$F^5(2C_2)$	$-(P)$	$= F^4(2C_2)$
	$-(d_1)$	$= F^4(3C_2)$
[6]	$-(b_1)$	$= F^4(B_3+C_2)$
	$-(a)$	$= F^4(B_4)$
	$-(d_1 d_2)$	$= F^4(4C_2)$
	$-(b_1 d_1)$	$= F^4(B_3+2C_2^*)$
$F^5(B_3)$	$-(P)$	$= F^4(B_3)$
	$-(d)$	$= F^4(B_3+C_2)$
[7]	$-(b')$	$= F^4(B_4^*)$
	$-(b_1)$	$= F^4(B_4)$
	$-(b'd)$	$= F^4(B_4^*+C_2)$
$F^5(B_3+C_2)$	$-(P)$	$= F^4(B_3+C_2)$
	$-(d)$	$= F^4(B_3+2C_2^*)$
[7]	$-(b)$	$= F^4(B_4^*+C_2)$
	$-(a)$	$= F^4(B_5)$
	$-(bd)$	$= F^4(B_4^*+2C_2^*)$
$F^5(B_4)$	$-(P)$	$= F^4(B_4)$
[7]	$-(b)$	$= F^4(B_5)$
	$-(e)$	$= F^4(U_6)$
$F^5(B_5)$	$-(P)$	$= F^4(B_5)$
[8]	$-(e)$	$= F^4(U_7)$

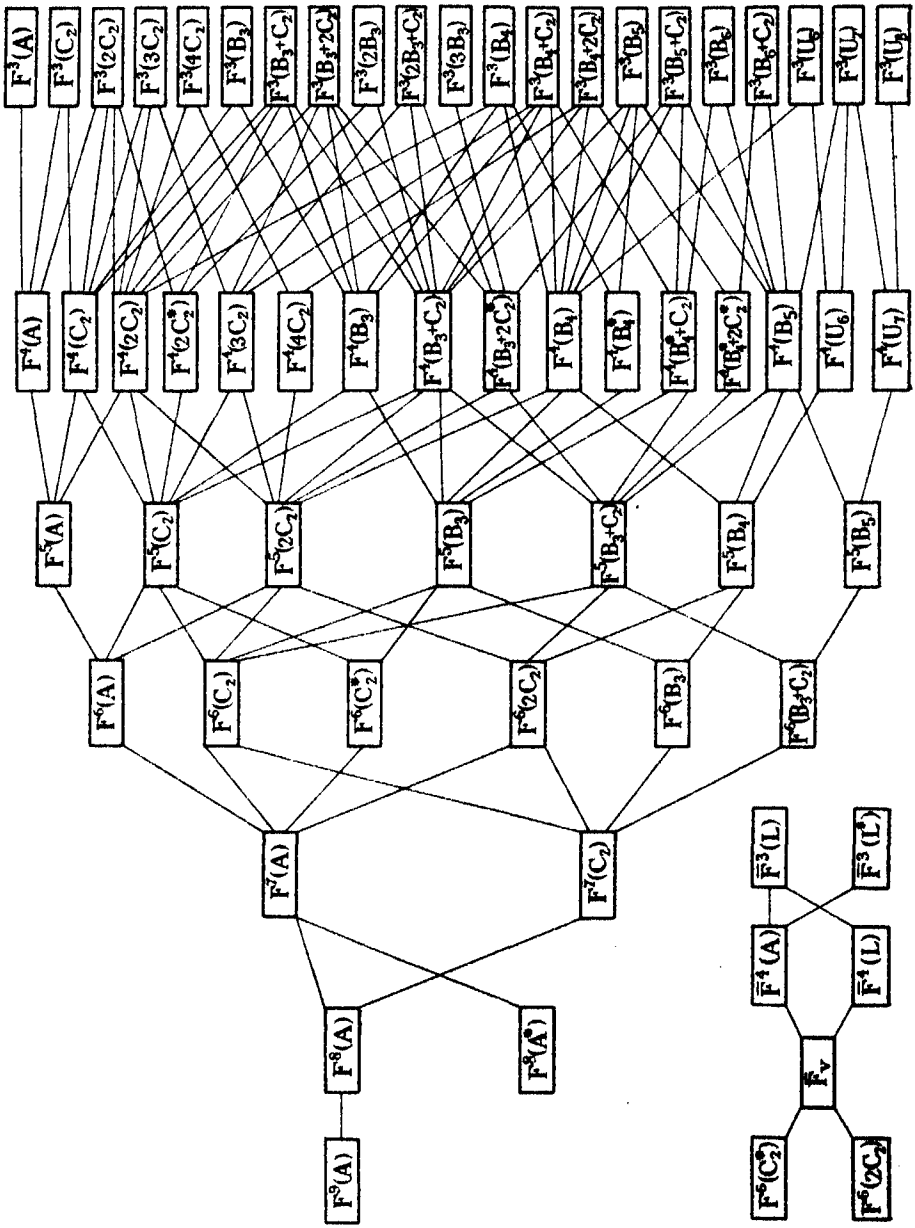
F^4 projected into F^3.

$F^4(A)$	$-(P)$	$=F^3(A)$
[4]	$-(1)$	$=F^3(C_2)$
	$-(1,12)$	$=F^3(2C_2)$
$F^4(C_2)$	$-(P)$	$=F^3(C_2)$
	$-(d_1)$	$=F^3(2C_2)$
[5]	$-(b_1)$	$=F^3(B_3)$
	$-(d_1d_2')$	$=F^3(3C_2)$
	$-(b_1d_1)$	$=F^3(B_3+C_2)$
$F^4(2C_2)$	$-(P)$	$=F^3(2C_2)$
	$-(13')$	$=F^3(3C_2)$
[5]	$-(b_1)$	$=F^3(B_3+C_2)$
	$-(a)$	$=F^3(B_4)$
	$-(13',24')$	$=F^3(4C_2)$
	$-(b_1,13')$	$=F^3(B_3+2C_2)$
$F^4(2C_2{}^*)$	$-(P)$	$=F^3(2C_2)$
[6]	$-(b_1)$	$=F^3(B_3+C_2)$
	$-(b_1b'_1)$	$=F^3(2B_3)$
$F^4(3C_2)$	$-(P)$	$=F^3(3C_2)$
	$-(b_1)$	$=F^3(B_3+2C_2)$
[6]	$-(a_1)$	$=F^3(B_4+C_2)$
	$-(b_1b_2')$	$=F^3(2B_3+C_2)$
$F^4(4C_2)$	$-(P)$	$=F^3(4C_2)$
[6]	$-(a_{12})$	$=F^3(B_4+2C_2)$
$F^4(B_3)$	$-(P)$	$=F^3(B_3)$
[6]	$-(d_1)$	$=F^3(B_3+C_2)$
	$-(b_1)$	$=F^3(B_4)$
	$-(d_1d_4')$	$=F^3(B_3+2C_2)$
	$-(b_1d_1)$	$=F^3(B_4+C_2)$
$F^4(B_3+C_2)$	$-(P)$	$=F^3(B_3+C_2)$
	$-(d_1)$	$=F^3(B_3+2C_2)$
	$-(b')$	$=F^3(2B_3)$
[6]	$-(b_1)$	$=F^3(B_4+C_2)$
	$-(a)$	$=F^3(B_5)$
	$-(b'd_1)$	$=F^3(2B_3+C_2)$
	$-(b_1d_1)$	$=F^3(B_4+2C_2)$
$F^4(B_3+2C_2{}^*)$	$-(P)$	$=F^3(B_3+2C_2)$
	$-(b_1)$	$=F^3(2B_3+C_2)$
[6]	$-(a_1)$	$=F^3(B_5+C_2)$
	$-(b_1b_2)$	$=F^3(3B_3)$
$F^4(B_4)$	$-(P)$	$=F^3(B_4)$
	$-(d_1)$	$=F^3(B_4+C_2)$
[6]	$-(b_1)$	$=F^3(B_5)$
	$-(e)$	$=F^3(U_6)$
	$-(d_1d_2)$	$=F^3(B_4+2C_2)$
	$-(b_1d_1)$	$=F^3(B_5+C_2)$
$F^4(B_4{}^*)$	$-(P)$	$=F^3(B_4)$
[7]	$-(b_1)$	$=F^3(B_5)$
$F^4(B_4{}^*+C_2)$	$-(P)$	$=F^3(B_4+C_2)$
[7]	$-(b_1)$	$=F^3(B_5+C_2)$
	$-(a)$	$=F^3(B_6)$
$F^4(B_4{}^*+2C_2{}^*)$	$-(P)$	$=F^3(B_4+2C_2)$
[7]	$-(a_1)$	$=F^3(B_6+C_2)$
$F^4(B_5)$	$-(P)$	$=F^3(B_5)$
	$-(d)$	$=F^3(B_5+C_2)$
[7]	$-(b)$	$=F^3(B_6)$
	$-(e)$	$=F^3(U_7)$
	$-(bd)$	$=F^3(B_6+C_2)$
$F^2(U_6)$	$-(P)$	$=F^3(U_6)$
[7]	$-(b_1)$	$=F^3(U_7)$
$F^4(U_7)$	$-(P)$	$=F^3(U_7)$
[8]	$-(b)$	$=F^3(U_8)$

F^3.

$F^3(A)$	[3]
(C_2)	[4]
$(2C_2)$	[4]
$(3C_2)$	[5]
$(4C_2)$	[5]
(B_3)	[5]
(B_3+C_2)	[5]
(B_3+2C_2)	[5]
$(2B_3)$	[6]
$(2B_3+C_2)$	[6]
$(3B_3)$	[6]
(B_4)	[5]
(B_4+C_2)	[6]
(B_4+2C_2)	[6]
(B_5)	[6]
(B_5+C_2)	[6]
(B_6)	[7]
(B_6+C_2)	[7]
(U_6)	[6]
(U_7)	[7]
(U_8)	[8]

The diagram expresses the facts of the table in genealogical style. The symbols indicating the position of the point of projection are omitted, for want of room.


$F^3(A)$
$F^3(C_2)$
$F^3(2C_2)$
$F^3(3C_2)$
$F^3(4C_2)$
$F^3(B_3)$
$F^3(B_3+C_2)$
$F^3(B_3+2C_2)$
$F^3(2B_3)$
$F^3(2B_3+C_2)$
$F^3(3B_3)$
$F^3(B_4)$
$F^3(B_4+C_2)$
$F^3(B_4+2C_2)$
$F^3(B_5)$
$F^3(B_5+C_2)$
$F^3(B_6)$
$F^3(B_6+C_2)$
$F^3(U_6)$
$F^3(U_7)$
$F^3(U_8)$
$F^4(A)$
$F^4(C_2)$
$F^4(2C_2)$
$F^4(2C_2^*)$
$F^4(3C_2)$
$F^4(4C_2)$
$F^4(B_3)$
$F^4(B_3+C_2)$
$F^4(B_3+2C_2^*)$
$F^4(B_4)$
$F^4(B_4^*)$
$F^4(B_4^*+C_2)$
$F^4(B_4^*+2C_2^*)$
$F^4(B_5)$
$F^4(U_6)$
$F^4(U_7)$
$F^5(A)$
$F^5(C_2)$
$F^5(2C_2)$
$F^5(B_3)$
$F^5(B_3+C_2)$
$F^5(B_4)$
$F^5(B_5)$
$F^6(A)$
$F^6(C_2)$
$F^6(C_2^*)$
$F^6(2C_2)$
$F^6(B_3)$
$F^6(B_3+C_2)$
$F^7(A)$
$F^7(C_2)$
$F^8(A)$
$F^8(A^*)$
$F^9(A)$
$\bar{F}^3(L)$
$\bar{F}^3(L^*)$
$\bar{F}^4(A)$
$\bar{F}^4(L)$
$\bar{F}_V$
$\bar{F}^6(C_2^*)$
$\bar{F}^6(2C_2)$

§ 9. *The Representation in the Plane.*†

9*a*. *Direct Application to the Generation of Nodal Surfaces by Projection.*—The modification introduced in the plane representation by the projection of a rational surface from a point is the subjection of the representative system to one linear condition ; in particular, if the point lies on the surface, this condition is to pass through the corresponding point of the plane.

Conversely, given the plane representation of a rational surface, it may thus be obtained as the projection of a non-singular surface, for, at the worst, the system of all curves of a given order represents a non-singular surface.

There is, however, occasionally some practical difficulty in choosing the space from which to make the projection, *e.g.*, in cases for which several of the base-points are adjacent.

A plane representation of any nodal cubic surface is readily obtained by projection from the node, the plane sections projecting into the cubics which pass through the six points (distinct or otherwise), in which the six rays meet the plane.

9*bi*. *Note.*—Multiple points of rational surfaces are represented in the plane in several ways.

(α) Isolated nodes.

- (i) By adjacent base-points,

  E.g., a node of a cubic surface other than that from which the projection is made to obtain the representation as suggested above.

- (ii) By a curve in the plane apparently representing a curve of zero order ; such always represents a single point, which, however, may be simple [*cf.* § 3*a*vi]. Several such curves may represent the same point, viz., if they intersect outside the base-points.

  E.g., the node of a cubic surface, from which the projection is made, represented by the projection of the nodal cone. In particular, for a binode, the two lines into which this degenerates.

  The join of the nodes in the representation of F^{3*} represents a simple point only.

- (iii) By a set of points such that all curves of the system through one pass through all. This is not possible for a surface in three dimensions.

- (iv) By combinations of the above.

(β) Multiple curves.

- (i) By a multiple base-point (such a point does not, of course, necessarily represent a multiple curve), *e.g.*, $\bar{F}^3$ (L), [§ 9*d*ii].

---

† *Cf.* Caporali, 'Memorie di Geometria,' p. 171.

(ii) By a set of curves between which the representative system determines an involutory correspondence, *i.e.*, each set of the correspondence, consisting of sub-sets on the various curves, is such that a representative curve through any one point of the set passes through all the others. The neighbourhood of a point is included as a particular case of a curve, its points being the directions of the tangents, *e.g.*, $\overline{\mathrm{F}}^{2}(\mathrm{L}^{*})$ [§ 9*d*iii]. In this way (ii) includes (i).

9*b*ii. The method may be extended to any rational variety, *e.g.*, Segre's cubic locus, represented by quadrics through five fixed points, whose ten joins represent the nodes.

9*c*. According to del Pezzo [*cf.* § 3*b*iv], a biplanar node on a cubic surface may be obtained by projecting an F^n from $n-3$ of the points in which an $\overline{n-1}$-secant $[n-3]$ meets the surface, the tangent planes at the two remaining points giving rise to the biplanes.

It appears, however, that *no* $\overline{k+2}$*-secant* $[k]$*s exist*, ($k \neq n-2$), *save those containing curves of the surface.* For consider, with the aid of the plane representation, what $\overline{k+2}$-secant $[k]$s are possible.

The $k+2$ intersections with the surface of a $\overline{k+2}$-secant $[k]$ do not present independent conditions to an $[n-1]$ (*i.e.*, a prime), which is to pass through them, and are therefore represented in the plane by $k+2$ points which do not present independent conditions to a cubic through $9-n$ points, *i.e.*, altogether $11-n+k$ points not independent with respect to a cubic.

A set of points not independent with respect to a cubic must be :—

5 or more points on a line, or
8 or more points on a conic, or
9 or more points on a cubic.

Suppose that of the set r points are base-points, then

Five points on a line of the plane lead to $5-r$ points of a $\overline{3-r}$-ic curve of the surface, so that any flat space through the $5-r$ points contains a $\overline{3-r}$-ic of the surface.

Eight points on a conic of the plane lead to $8-r$ points of a $\overline{6-r}$-ic of the surface, so that any flat space through the $8-r$ points contains a $\overline{6-r}$-ic of the surface.

Nine points on a cubic of the plane lead to $\overline{9-r}$ points not in general lying on an $\overline{8-r}$-ic of the surface; in this case, however, $11-n+k \geqslant 9$, *i.e.*, $k \geqslant n-2$.

Hence every $\overline{k+2}$-secant $[k]$ contains a curve (of order m, say) of the surface, except when $k=n-2$.

The surface is to be projected from $n-3$ points, *i.e.*, from an $[n-4]$, in the $[n-3]$; but the $[n-4]$ has m intersections with the m-ic, which therefore projects into an ordinary conic-node (excluding, of course, the special case in which the m-ic is degenerate).

9*d*. Note on the plane representation of projections of the Veronese surface.

It is known that—

$\overline{F}_V$ is represented by the conics of the plane.
$\overline{F}^4$(A) is represented by a general ∞^4 system of conics.
F^4(L) is represented by the ∞^4 system of conics apolar to a fixed pair of points.
Steiner surface is represented by an ∞^3 system of conics without base-points.
F^3(L) and $\overline{F}^3$(L*) are represented by an ∞^3 system of conics with one base-point.

An ∞^3 system of conics is outpolar to a pencil of class-conics, in particular to the three point-pairs, and so cuts in involution their three "axes." These represent the three double lines of the Steiner surface. Two (or three) coincide when two (or three) of the common tangents of the class pencil coincide. Two pairs of the common tangents coincide when the ∞^3 system has (their intersection as) a fixed base-point. Thus any ∞^3 system of conics with one base-point cuts a unique line in involution: this line represents the double line and the self-corresponding points the pinch-points; these coincide when the involution is parabolic; the line then passes through the base-point, and the system determines a (1 : 1) correspondence between the tangents at the point and the points of the line, which together represent the double line: the simple directrix, represented by the base-point, appropriately coincides with the double line.

9*d*ii. To a general ∞^3 system of conics with one base-point apply a quadratic transformation of which a vertex and opposite side are the base-point and line of involution, obtaining: a system of cubics with one node and two simple base-points; the nodal tangents form an involution.

9*d*iii. To a "parabolic" ∞^3 system of conics with one base-point apply a quadratic transformation with one vertex at the base-point, obtaining: a system of cubics with one node and two simple base-points, whose join is in (1 : 1) correspondence with a certain line through the node.

Apply a further transformation with vertices at the node and *one* of the simple base-points, obtaining: a system of cubics with a node, one fixed nodal tangent, and one simple base-point. The variable tangents at the base-points are in (1 : 1) correspondence.

§ 10. *Verification by Algebraical Methods.*

10*a*. *Equations to the Non-Singular* F^n.

F^9 (A), F^8 (A), F^7 (A), F^6 (A), represented in the plane by the cubics through 0, 1, 2, 3 general points respectively (which may be taken among the vertices of the triangle of reference), have parametric equations of the form

$$x_0 : x_1 : x_2 : x_3 : x_4 : x_5 : x_6 : x_7 : x_8 : x_9 = xyz : x^2y : x^2z : y^2z : y^2x : z^2x : z^2y : x^3 : y^3 : z^3,$$

the last 0, 1, 2, 3 co-ordinates being omitted in the respective cases.

If the products of all pairs of co-ordinates be considered, the identities between their expressions in terms of x, y, z lead to the following quadratic equations satisfied by all points upon the surfaces; it may be shown that they are not satisfied by any other points.

$$
\begin{array}{llll}
x_0^2 = x_1x_6 = x_2x_3 = x_4x_5 & & & \\
x_0x_1 = x_2x_4 = x_3x_7 & x_1^2 = x_4x_7 & x_1x_4 = x_7x_8 & x_4^2 = x_1x_8 \\
x_0x_2 = x_1x_5 = x_6x_7 & x_2^2 = x_5x_7 & x_2x_5 = x_7x_9 & x_5^2 = x_2x_9 \\
x_0x_3 = x_4x_6 = x_5x_8 & x_3^2 = x_6x_8 & x_3x_6 = x_8x_9 & x_6^2 = x_3x_9 \\
x_0x_4 = x_1x_3 = x_2x_8 & & & \\
x_0x_5 = x_2x_6 = x_1x_9 & & & \\
x_0x_6 = x_3x_5 = x_4x_9 & & & \\
x_0x_7 = x_1x_2 & & & \\
x_0x_8 = x_3x_4 & & & \\
x_0x_9 = x_5x_6 & & &
\end{array}
$$

F^5 (A).—Similarly by considering the cubics passing through (1, 1, 1) as well as through the vertices of the triangle of reference, the following six *sufficient* equations for the F^5 (A) are obtained :—

$$
\begin{aligned}
x_1x_3 + x_4x_6 = x_1x_6 = x_1x_5 + x_2x_6 \\
x_1x_3 + x_2x_4 = x_2x_3 = x_3x_5 + x_2x_6 \\
x_2x_4 + x_1x_5 = x_4x_5 = x_3x_5 + x_4x_6
\end{aligned}
$$

There is a linear syzygy between the three four-termed equations, so that 5 (= 6 − 1) quadratic equations suffice to determine the surface exclusively.

N.B.—Certain of the edges of the simplex of reference lie on the surfaces : they may be taken thus :—

Edge of the simplex of reference.	Notation for the lines.			
	F^5	F^6	F^7	F^8
(x_1x_2)	34	1		
(x_2x_5)	25	2		
(x_5x_6)	14	3	d_2	d
(x_6x_3)	35	4	d'	
(x_3x_4)	24	5	d_1	
(x_4x_1)	15	6		

F^4 (A) is the intersection of two general quadrics.

F^8 (A*) is represented by the quartics with two fixed nodes (say, the vertices (y), (z) of the triangle of reference), *i.e.*, expressed parametrically by

$$x_7 : x_2 : x_1 : x_5 : x_4 : x_0 : x_6 : x_3 : x_8 = y^2z^2 : xy^2z : xyz^2 : x^2y^2 : x^2z^2 : x^2yz : x^3y : x^3z : x^4,$$

which leads to the *sufficient* equations :—

$$\begin{array}{ll}
x_0{}^2 = x_1x_6 = x_2x_3 = x_4x_5 = x_7x_8 & x_1{}^2 = x_4x_7 \\
x_0x_1 = x_2x_4 = x_3x_7 & x_2{}^2 = x_5x_7 \\
x_0x_2 = x_1x_5 = x_6x_7 & x_3{}^2 = x_4x_8 \\
x_0x_3 = x_4x_6 = x_1x_8 & x_6{}^2 = x_5x_8 \\
x_0x_4 = x_1x_3 & \\
x_0x_5 = x_2x_6 & \\
x_0x_6 = x_3x_5 = x_2x_8 & \\
x_0x_7 = x_1x_2 & \\
x_0x_8 = x_3x_6 &
\end{array}$$

The equations involving only $x_0 \dots x_7$ are the same as those given for F^7 (A) and F^6 (A).

10*b*. *Note on the Analytical Method employed for Projection.*—Projection from an $[i]$ (*i.e.*, from $i+1$ independent points) on to an $[s]$ consists in finding those points of $[s]$ which are joined to $[i]$ by an $[i+1]$ meeting the variety to be projected.

Analytically, if $(\xi_0, \xi_1, \xi_2, \dots)$, $(\eta_0, \eta_1, \eta_2, \dots)$, $(\zeta_0, \zeta_1, \zeta_2, \dots) \dots$ be any $i+1$ independent points defining the $[i]$, any point of the $[i+1]$ which joins the $[i]$ to a point $x_0, x_1, \dots$ of $[s]$ is of the form

$$(x_0 + \lambda\xi_0 + \mu\eta_0 + \nu\zeta_0 + \dots,\ x_1 + \lambda\xi_1 + \mu\eta_1 + \dots,\ x_2 + \lambda\xi_2 + \mu\eta_2 + \dots,\ \dots,\ \dots),$$

where $\lambda, \mu, \nu \dots$ take any values.

Hence if a variety contains a point or points of the $[i+1]$, the $(\lambda, \mu, \nu \dots)$-eliminants of the equations obtained by substituting these co-ordinates in the equations of the variety will vanish, *i.e.*, these are the equations of the projected variety.

Remarks.

10*b*ii. It is usually convenient to take a component of the simplex of reference as the $[s]$: the only condition is that the $[s]$ must not meet the $[i]$.

10*b*iii. If $[i]$ is a component of the simplex of reference, this reduces to eliminating the appropriate co-ordinates.

10*b*iv. The case of F^4, *i.e.*, the intersection of two quadrics, $f = 0$, $g = 0$.

The projection from the point y is the λ-eliminant of

$$\begin{cases} f(x_0 + \lambda y_0, x_1 + \lambda y_1, \dots x_4 + \lambda y_4) = 0. \\ g(x_0 + \lambda y_0, x_1 + \lambda y_1, \dots x_4 + \lambda y_4) = 0. \end{cases}$$

i.e., of

$$\begin{cases} f_{xx} + 2\lambda f_{xy} + \lambda^2 f_{yy} = 0. \\ g_{xx} + 2\lambda g_{xy} + \lambda^2 g_{yy} = 0. \end{cases}$$

If y lies on the surface, $f_{yy} = 0 = g_{yy}$, and the λ-eliminant is

$$f_{xx}\, g_{xy} - f_{xy}\, g_{xx} = 0.$$

10*c*. The generation of each non-ruled F^3, by a projection (not in general unique) of the non-singular F^n in the fewest dimensioned space possible, will now be verified by algebraical methods.

10*ci*α. $F^3(C_2)$ and $F^3(2C_2)$ are the projections of the intersection of two quadrics in [4] from a point of a line thereon, and from the intersection of two lines thereon, respectively. Let the edges (x_0x_1), (x_0x_2) of the simplex of reference be two lines of the surface which is then expressed by

$$\left.\begin{array}{r} x_0u + x_1v + x_2w + x_1x_2 = 0 \\ x_0u' + x_1v' + x_2w' + x_1x_2 = 0 \end{array}\right\} \text{ where } u,\ v,\ w,\ u',\ v',\ w' \text{ do not involve } x_0,\ x_1,\ x_2.$$

$$\text{Eliminating } x_1 : \begin{vmatrix} x_0u + x_2w & v + x_2 \\ x_0u' + x_2w' & v' + x_2 \end{vmatrix} = 0, \text{ a cubic surface with a node at } (x_0).$$

$$\text{Eliminating } x_0 : \begin{vmatrix} u & x_1v + x_2w + x_1x_2 \\ u' & x_1v' + x_2w' + x_1x_2 \end{vmatrix} = 0, \text{ a cubic surface with nodes at } (x_1),\ (x_2).$$

10*ci*β. $F^3(3C_2)$ and $F^3(4C_2)$ *treated as explanatory examples.*

According to the table $F^3(4C_2) = F^4(2C_2) - (13', 24')$ and $F^4(2C_2) = F^5(A) - (15, 34)$, *i.e.*, to obtain $F^3(4C_2)$, $F^5(A)$ may be projected from the join of $(15, 34)$ and "$(13', 24')$," *i.e.*, the point of $F^5(A)$ which projects into $(13', 24')$ of $F^4(2C_2)$, viz., $(45, 13)$. $(45, 13)$ may for convenience be replaced by $(35, 14)$, a vertex of the simplex of reference, for the pairs of points $(15, 34)$, $(45, 13)$; $(15, 34)$, $(35, 14)$ are precisely alike. Symbolically $F^3(4C_2) = F^4(2C_2) - (13', 24') = F^5(A) - (15, 34) - (13', 24') = F^5(A) - \overline{(15, 34) + (45, 13)} = F^5(A) - \overline{(15, 34) + (35, 14)}$. $F^3(4C_2)$ is therefore obtained by projecting $F^5(A)$ from the vertices (x_1), (x_6) of the simplex of reference; eliminating x_1, x_6 from the equations to $F^5(A)$:—

$$x_2(x_3x_5 - x_4x_5) = x_4(x_3x_5 - x_2x_3),$$

i.e., $-\dfrac{1}{x_2} - \dfrac{1}{x_3} + \dfrac{1}{x_4} + \dfrac{1}{x_5} = 0$, the standard equation of a four-nodal cubic.

An algebraically more general process of projection is exemplified by the case of $F^3(3C_2) = F^5(A) - \overline{(34) + (35, 14)}$, *i.e.*, project from the vertex (x_6) and any point of the edge $(x_1\, x_2)$.

$$\left.\begin{array}{ll}
\left.\begin{array}{l}[i]\text{, the line-vertex of projection}\\ \quad\text{is defined by the points}\end{array}\right\} & \left\{\begin{array}{ccccccc}1&\phi&0&0&0&0&\lambda\\ 0&0&0&0&0&1&\mu\end{array}\right.\\
\left.\begin{array}{l}\text{A point of }[s]\text{, a chosen solid not}\\ \quad\text{meeting }[i]\text{ is}\end{array}\right\} & \begin{array}{ccccccc}&x_2&x_3&x_4&x_5&&1\end{array}
\end{array}\right\}\begin{array}{c}\lambda,\ \mu,\ 1\text{ are}\\ \text{arbitrary}\\ \text{multipliers.}\end{array}$$

If their join meets $F^5(A)$ $\left\{\begin{array}{l} x_3x_5 + x_4\mu = x_4x_5. \\ \lambda x_3 + x_4\mu = \lambda\mu. \\ x_2x_4 + \lambda x_5 = x_4x_5. \end{array}\right.$

Eliminating λ, μ : $(1-\phi)\, x_3x_4x_5 + x_2x_4x_5 - x_2x_3x_5 - x_2x_3x_4 + \phi x_4^2x_5 = 0$, which is a cubic surface with nodes at (x_2), (x_3), (x_5).

The successive projections are usually most easily combined into one by inspection of the diagrams. Different successions sometimes form the same combined projection.

Surfaces with B_3s.

10ciiα.

$$F^3(B_3) = F^6(A) - \overline{(15) + (34)}$$

$$F^3(B_3 + C_2) = F^6(A) - \overline{(15) + (25, 34)}$$

$$F^3(B_3 + 2C_2) = F^6(A) - \overline{(15, 24) + (25, 34)}.$$

The line-vertex of projection is
(where, for $F^3(B_3 + C_2)$, $\phi = \infty$;
for $F^3(B_3 + 2C_2)$, $\phi = \psi = \infty$)
Point of arbitrary solid

$$\begin{array}{lcccccccc} & 1 & \phi & 0 & 0 & 0 & 0 & & \lambda \\ & 1 & 0 & 0 & \psi & 0 & 0 & & \mu. \\ \text{Point of arbitrary solid} & x_1 & & x_3 & & x_5 & x_6. & & \end{array}$$

If their join meets $F^6(A)$ $\left\{\begin{array}{r} \phi\lambda\,(x_6 - x_3) + x_3x_5 = 0 \\ \psi\mu\,(x_6 - x_3) + x_3x_5 = 0 \\ (x_1 + \lambda + \mu)\,(x_3 - x_6) + \psi\mu x_6 = 0. \end{array}\right.$

Eliminating λ, μ : $x_1(x_6 - x_3)(x_6 - x_5) - x_3x_5\left(\dfrac{x_6 - x_5}{\phi} - \dfrac{x_6 - x_3}{\psi}\right) + x_3x_5x_6 = 0.$

Binode at (x_1) ; biplanes $(x_5 - x_6)(x_3 - x_6) = 0$.

Put $\phi = \infty$. Additional node at (x_5) ; nodal cone $-x_1(x_6 - x_3) - \dfrac{x_3}{\psi}(x_6 - x_3) + x_3x_6 = 0$.

Put $\phi = \psi = \infty$. Additional node at (x_3) ; nodal cone $-x_1(x_6 - x_5) + x_5x_6 = 0$.

10ciiβ.

$$F^3(2B_3) = F^6(A) - \overline{(5, 6) + (1) + (4)}$$

$$F^3(2B_3 + C_2) = F^6(A) - \overline{(5, 6) + (1) + (3, 4)}$$

$$F^3(3B_3) = F^6(A) - \overline{(5, 6) + (1, 2) + (3, 4)}.$$

The plane-vertex of projection is
(where, for $F^3(2B_3 + C_2)$, $\phi = \infty$;
for $F^3(3B_3)$, $\phi = \psi = \infty$)

$$\begin{array}{lcccccccc} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \\ & 0 & 1 & \psi & 0 & 0 & 0 & 0 & \lambda \\ & 0 & 0 & 0 & 1 & 0 & 0 & \phi & \mu. \\ \text{Point of arbitrary solid} & x_0 & x_1 & & x_3 & & x_5. & & \end{array}$$

If their join meets $F^6(A)$ $\left\{\begin{array}{l} x_0\lambda\psi = (x_1 + \lambda)\,x_5 \\ x_0x_5 = \psi\lambda\phi\mu \\ x_0\mu\phi = (x_3 + \mu)\,x_3. \end{array}\right.$

Eliminating λ, μ : $x_0(x_0\psi - x_3)(x_0\phi - x_5) = \phi\psi x_1x_3x_5$.

Binodes at $\left\{\begin{array}{l} (x_1)\text{; biplanes } x_3x_5 = 0 \\ (x_3)\text{; biplanes } x_1x_5 = 0. \end{array}\right.$

Put $\phi = \infty$: $x_0^2(x_0\psi - x_5) = \psi x_1x_3x_5$. Additional node at (x_5) ; nodal cone $x_0^2 + \psi x_1x_3$.
Put also $\psi = \infty$: $x_0^3 = x_1x_3x_5$. The conic-node at (x_3) becomes a binode.

Surfaces with a B_4.

10ciiiα.

$$F^3(B_4) = F^4(4C_2) - (a) = F^5(A) - (\text{tangent line at } (15, 34)).$$

The tangent plane at (15, 34) is $x_3 = x_5 = x_6 = 0$.

Thus the line-vertex of projection is $\left\{\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & & \\ 0 & 1 & 0 & \phi & 0 & 0 & & \lambda. \end{array}\right.$

Point of arbitrary solid $\quad x_3 \quad x_4 \quad x_5 \quad x_6.$

If their join meets $F^5(A)$ $\left\{\begin{array}{r} \lambda(x_6 - x_3) + x_3x_5 = 0 \\ x_3x_5 + (x_4 + \lambda\phi)(x_6 - x_5) = 0. \end{array}\right.$

Eliminating λ: $x_4(x_6 - x_5)(x_6 - x_3) + x_3x_5\{x_6 - x_3 - \phi(x_6 - x_5)\} = 0.$

Binode B_4 at (x_4); biplanes $(x_6 - x_3)(x_6 - x_5) = 0$.

10ciiiβ.

$$F^3(B_4 + C_2) = F^6(A) - \overline{(3) + (4) + (5, 6)}$$

$$F^3(B_4 + 2C_2) = F^6(A) - \overline{(2, 3) + (4) + (5, 6)}$$

The plane-vertex of projection is (where for $F^3(B_4 + 2C_2)$, $\psi = 0$) $\left\{\begin{array}{ccccccccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 & & \\ 0 & 0 & 0 & 1 & 0 & 0 & \phi & & \lambda \\ 0 & 0 & 0 & 0 & 0 & 1 & \psi & & \mu \end{array}\right.$

Point of arbitrary solid $\quad x_0 \quad x_1 \quad x_2 \quad x_3.$

If their join meets $F^6(A)$ $\left\{\begin{array}{l} x_0x_2 = x_1\mu \\ x_0^2 = x_2(x_3 + \lambda) \\ x_0^2 = x_1(\lambda\phi + \mu\psi). \end{array}\right.$

Eliminating λ, μ: $x_0^2x_2 = \phi x_0^2x_1 - \phi x_1x_2x_3 + \psi x_0x_2^2.$

Binode B_4 at (x_3); biplanes $x_1x_2 = 0$.

Conic-node at (x_1); nodal cone $x_0^2 - x_2x_3 = 0$.

Put $\psi = 0$. Additional conic-node at (x_2); nodal cone $x_0^2 + \phi x_1x_3 = 0$.

Surfaces with a B_5.

10civ.

$$F^3(B_5) = F^6(A) - \overline{(\text{Tangent line at } (5, 6)) + (4)}.$$

$$F^3(B_5 + C_2) = F^6(A) - \overline{(\text{Tangent line at } (5, 6)) + (3, 4)}.$$

The tangent plane at (5, 6) is $x_0 = x_2 = x_5 = x_6 = 0$.

The plane-vertex of projection is (where, for $F^3(B_5 + C_2)$, $\phi = \infty$) $\left\{\begin{array}{ccccccccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 & & \\ 0 & \psi & 0 & 1 & 0 & 0 & 0 & & \lambda \\ 0 & 0 & 0 & 1 & 0 & 0 & \phi & & \mu \end{array}\right.$

Point of arbitrary solid $\quad x_0 \quad x_1 \quad x_2 \qquad\qquad x_5.$

If their join meets $F^6(A)$ $\left\{\begin{array}{l} x_0x_2 = (x_1 + \lambda\psi)\,x_5. \\ x_0x_5 = x_2\phi\mu. \\ x_0\phi\mu = (\lambda + \mu)\,x_5. \end{array}\right.$

Eliminating λ, μ: $x_0x_2^2 - x_1x_2x_5 - \psi x_0^2x_5 + \dfrac{\psi}{\phi}x_0x_5^2 = 0.$

Binode B_5 at (x_1); biplanes $x_2x_5 = 0$. $\quad x_5 = 0$ meets the surface in $x_0x_2^2 = 0$.

Put $\phi = \infty$: $x_0x_2^2 - x_1x_2x_5 - \psi x_0^2x_5 = 0.$

Additional (conic-) node at (x_5); nodal cone $x_1x_2 + \psi x_0^2 = 0$.

Surfaces with a B_6.

10cv.

$$F^3(B_6) = F^7(A) - \overline{(\text{Tangent line at } (d'd_2)) + (\text{Tangent line at point of } d_1)}$$

$$F^3(B_6 + C_2) = F^7(A) - \overline{(\text{Tangent line at } (d'd_2)) + (\text{Tangent line at point of } d_1 \text{ to a conic of } F^7(A))}.$$

The tangent plane at $d'd_2$ is $x_0 = x_1 = x_2 = x_4 = x_7 = 0$.

The tangent plane at a point of d_1 (the co-ordinate-vertex (x_4) is sufficiently general) is

$$x_0 = x_2 = x_5 = x_6 = x_7 = 0.$$

The solid-vertex of projection is (where, for $F^3(B_6 + C_2)$, $\psi = 0$)

$$\left\{ \begin{array}{cccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & & \\ 0 & 0 & 0 & 1 & 0 & \phi & 0 & 0 & & \lambda \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & & \\ 0 & 1 & 0 & \psi & 0 & 0 & 0 & 0 & & \mu \end{array} \right.$$

Point of arbitrary solid $\quad x_0 \quad\quad x_2 \quad x_3 \quad\quad\quad\quad x_7$.

If their join meets F^7 (A)
$$\left\{ \begin{aligned} x_0^2 &= x_2(\mu\psi + \lambda + x_3) \\ x_0\mu &= x_7(\mu\psi + \lambda + x_3) \\ x_0x_2 &= \lambda\mu\phi. \end{aligned} \right.$$

Eliminating λ, μ : $x_2^3 = \phi x_7(x_0^2 - x_2x_3 - \psi x_0x_7)$.

Binode B_6 at (x_3) ; biplanes $x_2x_7 = 0$. $x_7 = 0$ meets the surface in $x_2^3 = 0$.

Put $\psi = 0$: $x_2^3 = \phi x_7(x_0^2 - x_2x_3)$.

Additional (conic-) node at (x_7).

Surfaces with Unodes.

10cvi.

$F^3(U_6) = F^6(A) -$ (The plane containing three consecutive points near (3, 4)).

Note.—Consecutive points ($\geqslant 3$ in number) must be distinguished from merely adjacent points. Through a set of consecutive points there pass curves which are simple in the neighbourhood of the set. Such a set does *not* define the tangent plane of a surface on which it lies.

The co-ordinates of the three points required are readily obtained with the aid of the plane representation ; through the three points passes just one quartic of the surface, to which corresponds in the plane a conic touching $x = 0$ at (z) and passing through (x). The conic can be taken in the parametric form $t^2, t, 1 - t$. By substituting in the parametric expressions for the surface, the quartic is obtained as

$$t^2 - t^3, t^4, t^3 - t^4, t - t^3, t^3, t - 2t^2 + t^3, 1 - 2t + t^2.$$

Thus three consecutive points at (d_3d_4) (*i.e.*, $t = 0$) are given by

$$\left\{ \begin{array}{ccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & t & 0 & t & 1 - 2t \\ t^2 & 0 & 0 & t - t^3 & 0 & t - 2t^2 & 1 - 2t + t^2 \end{array} \right. \quad \text{where } t \to 0.$$

These define the same plane-vertex of projection as

$$\left\{\begin{array}{ccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & & \lambda \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & & \mu. \end{array}\right.$$

Point of arbitrary solid $x_1 \quad x_2 \quad x_3 \quad x_4.$

If their join meets $F^6(A)$ $\left\{\begin{array}{r} x_2(x_3 + \lambda + \mu) = x_4\lambda \\ \mu x_1 = x_2x_4 \\ \mu x_3 = x_1\lambda. \end{array}\right.$

Eliminating λ, μ: $x_3x_1^2 = x_3x_4^2 - x_2^2x_4 - x_1x_2x_4.$

Unode at (x_3); uniplane $x_1 = 0$, meeting the surface in $x_2x_4(x_2 - x_4) = 0.$

10cvii.

$F^3(U_7) = F^7(A)$ — (Solid containing four consecutive points near $(d'd_2)$).

Through the four points passes a rational quintic curve, represented in the plane by t^3, t, 1; and hence parametrically expressed by $t^8, t^6, t^4, t^3, t^4, t^2, t, t^6$. The method employed for $F^3(U_8)$ may be applied.

The solid-vertex of projection is $\left\{\begin{array}{ccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & \lambda \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & \mu. \end{array}\right.$

Point of arbitrary solid $\quad x_1 \quad x_2 \quad x_3 \qquad\qquad x_7.$

If their join meets $F^7(A)$ $\left\{\begin{array}{r} (x_2 + \mu)\mu = (x_3 + \lambda)x_7. \\ x_1^3 = \mu x_7. \\ (x_2 + \mu)^2 = \lambda x_7. \end{array}\right.$

Eliminating λ, μ: $x_2^2x_7 + x_1^2x_2 + x_3x_7^2 = 0.$

Unode U_7 at (x_3); uniplane $x_7 = 0$, which meets the surface in $x_1^2x_2 = 0.$

10cviii.

$F^3(U_8) = F^6(A)$ — (A [5] containing six consecutive points of the surface, of which three lie on a cubic).

This is, of course, equivalent to generating $F^3(U_8)$ as a projection of $F^9(A)$.

Through the six points pass ∞^3 9-ics of the surface, any one of which may be represented in the plane by the cubic t^3, t, 1; and thus itself be expressed as $t^4, t^7, t^6, t^3, t^5, t^8, t, t^9, t^2, 1$, the consecutive points being near $t = 0$; they define the same [5] as the vertices (x_0), (x_2), (x_4), (x_6), (x_9) and the point (0, 0, 0, 0, 0, 1, 0, 0, 1, 0).

Projecting from the first five points: $x_2^2 = x_5x_7$, $x_1^3 = x_7^2x_8$, etc.

If the join of the sixth point $\quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \qquad \lambda,$

and $\quad x_1 \quad x_2 \qquad\qquad x_5 \qquad x_7,$

meets the surface so obtained $\left\{\begin{array}{l} x_2^2 = (x_5 + \lambda)x_8 \\ x_1^3 = x_7^2\lambda. \end{array}\right.$

Eliminating λ: $x_2^2x_7 = x_5x_7^2 + x_1^3.$

Unode U_8 at (x_5); uniplane $x_7 = 0$, meeting the surface in $x_1^3 = 0.$

The Spark Spectrum of Copper.

By A. C. MENZIES, M.A., Head of the Physics Department, University College, Leicester.

(Communicated by Prof. R. Whiddington, F.R.S.—Received March 3, 1928.)

The following is an account (A) of a determination of the lowest term of the spectrum of ionised copper (Cu II) and (B) of the testing of Shenstone's allocation of terms in this spectrum. While doing this some new terms were found (C).

(A) *Discussion of Sources.*

A method of obtaining spectrograms with one very short exposure (of the order of 1/100th second) by the fusing of wires has been described in a previous paper* by the author. When the wires are fused in air at atmospheric pressure, lines involving low energy levels are absorbed, while those involving outer levels tend to be displaced.

This method has now been used in the Schumann region, where it has the following advantages over the more normal arc and spark discharges in vacuo, which appear to have been used exclusively up to the present.

(1) Arc discharges are difficult to maintain, and this difficulty does not arise in the case of the fuse.

(2) The pressure of the gas in the spectrograph and around the source must be kept low while the exposure is being made, and the evolution of gases from the source raises the pressure. But when a fuse is used, the pressure need only be reduced preliminarily. When the fuse occurs, the light will have travelled to the grating and back to the plate before the gases from the fuse have reached the slit. In consequence, slow-acting pumps can be used, since low pressure has not to be maintained.

(3) Material shot out from the arc fouls the slit. This occurs also with the fuse, but not before the light has passed through the slit.

(4) With arcs and sparks, trouble is caused by "wandering" of the source, *i.e.*, during the protracted exposures required, the position of the source varies, so causing broadening of the lines. The fuse, on the other hand, is localised, and usually gives very sharp lines.

A Possible Disadvantage.—In the fuses in air previously investigated, displacements and reversals were frequent. It was therefore feared that the displacements possible might vitiate the measurement of wave-lengths in the Schumann region, when the fuse was the source employed. But there are two important differences in the present case: the medium is changed from air to vacuo, and much shorter wave-lengths are in question.

---

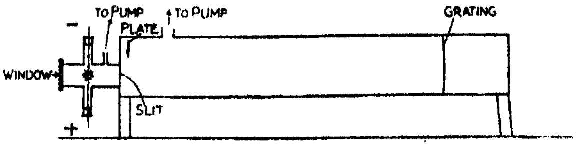
* 'Roy. Soc. Proc.,' A, vol. 117, p. 88 (1927).

To test this point, spectrograms were taken in the visible and ultra-violet of copper wires fused in vacuo, and compared with those of the copper arc (4 amperès) in air. The lines in the case of the fuse were very much sharper, and there were no reversals, not even λλ 3274 or 3247, although these two wave-lengths were very slightly symmetrically broadened. (This fits in with the observation previously made that broadening is the first step towards reversal.) No displacements could be observed, so that one can probably rely on measurements made in the Schumann region, using this method.

Also, when the fuses were made in air, displacements were most prevalent in the regions of longer wave-length (above about 3000 Å.U.), as is usual with pressure shifts. There were a few lines which showed a tendency to shift, of wave-length even as short as 2100 Å.U., but they were very rare. So shifts in the Schumann region would probably be still more unlikely.

(Results similar to those got by fuses in vacuo were obtained when wires were fused in an atmosphere of hydrogen. This agrees with the well-known "sharpening" of spectral-lines from an arc by running it in vacuo or in hydrogen. When the fuses were made under water, however, reversals were still present, though not so prominently as when the medium was air. The experiments with hydrogen are important since the vacuum grating spectroscope was in some cases washed out preliminarily with hydrogen, before evacuation.)

Apparatus Employed.—A cruciform chamber (see diagram) was waxed on to the slit-carrying recess of a Hilger vacuum grating spectrograph, having a concave N.P.L. grating of 1 metre radius, ruled with a grating interval of 0·0000694158 cms.



The main-tube of the fuse-chamber was of 1-inch brass tubing, and was closed at the end remote from the spectrograph with a glass window, to facilitate lining-up. Two cross-tubes of ⅜-inch brass tubing were mounted in line with one another, and were coned-out at their outer ends so as to take two tapered ebonite plugs. Each of these was bored with a small hole, excentrically, through which the wire might be thrust. When greased with rubber-compound, they were quite air-tight, and by twisting them the wire could be moved laterally. Exhaustion was carried out by a side-tube in the fuse-chamber, as well as by the port in the spectrograph. Only a small-size Cenco-Hyvac pump was available, in series with a two-stage glass mercury vapour pump. Exhaustion was carried on until no discharge would pass in a small discharge-tube attached to the spectrograph. With such a slow-acting pumping system, the slightest trace of a leak was fatal.

Procedure.—Two pieces of copper wire (20 gauge) were thrust through the stoppers, one in each, and these were then replaced and twisted so that the two ends met opposite the middle of the slit, and butted together. Their outer ends were joined to leads from a heavy double-bladed spring switch in series with an 110-volt, 144 ampère-hour accumulator-battery.

The apparatus was next washed out with dried hydrogen and exhausted as far as the pumps would allow.

On releasing the switch, so that it closed, the wires were fused at the join, so that a current of the order of 100 ampères passed momentarily. One such fuse was found to be sufficient to give a photograph when using quite a fine slit (*e.g.*, 0·1 mm.).

Results.—In the present work, the result of chief interest is the possible identification of the three lines $\mathrm{I}\,{}^1\mathrm{S}_0 - a\,{}^1\mathrm{P}_1$, $\mathrm{I}\,{}^1\mathrm{S}_0 - a\,{}^3\mathrm{D}_1{}^1$, and $\mathrm{I}\,{}^1\mathrm{S}_0 - a\,{}^3\mathrm{P}_1$, thus identifying the term $\mathrm{I}\,{}^1\mathrm{S}_0$, the basic term corresponding to the normal orbit of Cu II.

The allocation of the three outer orbit terms above is due to A. G. Shenstone,* who has ordered the terms of the copper spark spectrum, with the exception of this term $\mathrm{I}\,{}^1\mathrm{S}_0$, the ground-term required by the Heisenberg-Hund scheme. This term should combine with the three terms already mentioned to give three lines, which Shenstone estimated would probably be found to be within 100 Å.U. of the wave-lengths 1353, 1363 and 1466 Å.U. respectively. If my identification is correct, Shenstone's estimate is within 5 Å.U. !

Calculation of Wave-lengths.—In all spectra in the Schumann region, lines due to carbon and other impurities appear. These have been taken as standard lines, and the copper wave-lengths calculated by interpolation. The wave-lengths of some of these frequently occurring impurity lines have been determined with great accuracy by Smith and Lang,* and their values have been adopted here. The lines actually used were the two carbon lines 1561·46, 1335·70, and the Lyman hydrogen line 1215·68. Owing to the curvature given to the plate, the dispersion is nearly normal. Between 1216 and 1336 the dispersion was 17·526 and between 1336 and 1561 it was 17·530 Å.U. per millimetre.

---

* 'Phys. Rev.,' vol. 29, p. 380 (1927).

This leads to the values λ 1358·76, ν 73596·5 ; λ 1367·92, ν 73103·7 ; and λ 1472·38, ν 67917·4 for the three lines. The first line is believed to be the most accurately determined, owing to its strength and its proximity to one of the fiducial lines. The error in its determination is believed to be less than 0·05 Å.U. The other lines are probably not much more in error ; owing to their sharpness they were very easy to set upon with the microscope, but the third line in particular was well removed from the standard lines, so making the calculation of its wave-length more hazardous.

The differences of the wave-numbers of the lines should give the differences of the terms involved, since the lower level term is common to all. For convenience of comparison they are here set out :—

1st line—2nd line	492·8	$a\,{}^3D_1{}^1 - a\,{}^1P_1$	493·8
2nd line—3rd line	5186·3	$a\,{}^3P_1 - a\,{}^3D_1{}^1$	5185·5
1st line—3rd line	5679·1	$a\,{}^3P_1 - a\,{}^1P_1$	5679·3

The present identification of these three lines depends upon these wave-number differences, upon the relative intensities of the lines, and upon the nearness to the position estimated by Shenstone. Also 1358·76 is the strongest line on plates taken with this source (with the exception of the carbon lines at 1561) and one would expect $I\,{}^1S_0 - a\,{}^1P_1$ to be extremely strong in the fuse spectrum since this source has been found to emit low-level spark-lines in great intensity. Although $a\,{}^3P_1$ is a lower level than $a\,{}^1P_1$, $I\,{}^1S_0 - a\,{}^3P_1$ will be less strong, since it involves a transition from a triplet level to a singlet level, and such combination lines are usually weak.

If we take Shenstone's value 51667·1 for $a\,{}^1P_1$, then the term $I\,{}^1S_0$ is −21929·4, a negative term, since $a\,{}^3D_3$ has been up to the present the lowest known term, and has consequently been taken as zero.

(B) *An Attempt to Test Shenstone's Scheme.*

A. G. Shenstone† has given a consistent scheme of terms to describe the spectrum of the copper spark, involving wave-lengths lying between the limits

---

* 'Phys. Rev.,' vol. 27, p. 36 (1926).

† *Loc. cit.*

2900 and 1944 Å.U. But below λ 2356·65 Shenstone gives calculated values of the wave-lengths, which are based on the values of the terms obtained from lines occurring in a more accurately measured region of the spectrum. The lines so calculated are in every case near to lines previously measured, and differ from them by steadily increasing amounts.

J. C. McLennan and A. B. McLay* have recently ordered the spectrum of the gold arc (Au I) into terms, and some of the lines whose terms are known occur conveniently for comparison with some of Shenstone's copper spark lines. This is particularly true in the region where the calculated values of the copper lines deviate most from wave-lengths previously measured.

Thus it is possible to compare the wave-lengths of the gold lines, taking the calculated copper lines as standards, with the calculated values based on the term allocation of McLennan and McLay. If these are in agreement, then it must mean that the copper and the gold terms are both correctly assigned, or that both sets are wrong to the same extent and in the same direction—a very unlikely case.

Apparatus.—The source was a Pfund arc, operated at 3½ ampères from an accumulator battery of 110 volts, the light being emitted by a glowing bead of "gold"—roughly equal parts by weight of copper and gold. In this way the copper and gold lines to be compared were obtained side by side in the same spectrum, so that no difficulties owing to the shift of a comparison spectrum could arise.

The spectrograph was a small quartz-prism instrument made by Adam Hilger, Ltd. (E.37); with Schumann plates it will record down to 1850 Å.U., and has in the region about 2000 Å.U. a dispersion of about 12 Å.U. per millimetre. The error in measurement was probably less than 0·05 Å.U. Schumann plates were used.

Calculation of Wave-lengths.—The difficulty arises at the outset that the dispersion of air is not known below 2000 Å.U., but the correction from wave-lengths in air to those in vacuo decreases so regularly, that it might be extrapolated without much error. On the other hand, the air has already begun to absorb, and so the regularity of the correction-tables may be deceptive. Consequently the unorthodox procedure has here been adopted of measuring wave-lengths over this short region (2000 to 1944 Å.U.) directly in vacuum values; that is to say, the correction from air to vacuum values has been virtually included in the Hartmann formula employed. The justification for this is that the correction will presumably vary regularly through the region, and that in any case the region is so small that the change in the correction is bound to be small.

---

* 'Roy. Soc. Proc.,' A, vol. 112, p. 95 (1926).

Results.—Taking as standards λλ 2000·348, 1979·971, and 1944·606, which are three calculated copper lines, the following values of wave-lengths in the copper-gold spectrum have been measured, and are here compared with their values as calculated by Shenstone and by McLennan and McLay respectively.

Table I.

Wave-length (in vacuo).	Element and Terms.	Wave-number observed.	Wave-number calculated.
2022·02	Au $\mathrm{I}\,{}^2\mathrm{D}_3 - b_{23}$	49455·5	49455·4
2017·55	Cu $a\,{}^3\mathrm{D}_3 - a\,{}^3\mathrm{D}_2{}^1$	49565·1	49564·9
2016·21	Cu $a\,{}^3\mathrm{D}_1 - a\,{}^1\mathrm{P}_1$	49598·0	49597·4
2012·67	Au —	49685·2	—
2001·30	Au —	49967·5	—
2000·348	Cu $a\,{}^3\mathrm{D}_3 - a\,{}^1\mathrm{F}_3$	Standard	49991·3
1990·43 (ft.)	Au —	50240·4	—
1989·87	Cu $a\,{}^3\mathrm{D}_2 - a\,{}^3\mathrm{D}_1{}^1$	50254·5	50254·8
1984·17 (ft.)	Au —	50398·9	—
1979·971	Cu $a\,{}^3\mathrm{D}_2 - a\,{}^1\mathrm{D}_2{}^1$	Standard	50505·8
1978·07	Au —	50554·3	—
1970·51	Cu $a\,{}^3\mathrm{D}_2 - a\,{}^1\mathrm{P}_1$	50748·3	50748·6
1951·98	Au $\mathrm{I}\,{}^2\mathrm{S}_1 - 2\,{}^2\mathrm{P}_2$	51230·0	51231·0
1944·606	Cu $a\,{}^3\mathrm{D}_3 - a\,{}^1\mathrm{D}_2{}^1$	Standard,	51424·3
1939·21 (ft.)	Au $\mathrm{I}\,{}^2\mathrm{D}_3 - 3\,{}^2\mathrm{P}_2$	51567·4	51567·4
1919·57	Au $\mathrm{I}\,{}^2\mathrm{D}_3 - c_{23}$	52095·0	52094·6
1904·8 (approx.)	Au —	52500	—
1899·9 ,,	Au —	52630	—
1890·5 ,,	Au —	52900	—
1879·8 ,,	Au $\mathrm{I}\,{}^2\mathrm{S}_1 - 2\,{}^2\mathrm{P}_1$	53200	53208·3

It will be seen that the copper and gold wave-lengths are consistent with one another, so that most probably Shenstone's scheme for the copper spark and McLennan and McLay's scheme for the gold arc are both correct.

Further Confirmatory Evidence.—In order to see what further reversals would be occasioned by still more increasing the current through a fuse, some 12 gauge rods of copper were fused by causing them to short-circuit a 1000 k.w. transformer. (This was arranged through the courtesy of the Leicester Mains Department, during some fuse-testing.) The spectrograms so obtained were not much different from those produced by shorting the accumulator-battery. A few additional lines were observable in the far ultra-violet (c. 1850) and some additional lines showed self-reversal.

These made it possible to pick out the low-lying terms, and the result is to confirm Shenstone's choice. Practically all the lines which have as final orbits the lowest levels :—$a\,{}^3\mathrm{D}_{3,\,2,\,1}$ or $a\,{}^1\mathrm{D}_2$ are either absorbed or self-reversed.

The sharpness of the self-reversal made it possible to measure the wave-lengths of these lines with accuracy, and the calculated wave-lengths agreed well with the measured values.

(C) *Additional Terms.*

In the course of this work some additional higher terms were found in the copper spark spectrum. Without Zeeman effect observations it is practically impossible to decide on the exact nature of the terms, but their values are here given in case they may be of use to other workers in this field.

In the table which follows the wave-lengths are taken from Kayser's "Spektroskopie," vol. 7, and Hasbach's values have been adopted wherever possible. Where these have been wanting, other values have been used, and the authority is indicated by the letters standing first, in the second column, which also gives the intensities. H stands for Hasbach, E for Eder, EH for Exner and Haschek, A for Aretz, W for Wagner, and M for measurements made by the author.

Table II.

Wave-length.	Authority and Intensity.	Wave-number.	Terms.
4910·325	A, 2u ; EH, 2u	20359·6	$a\,{}^1P_1$ —72026·7
4851·8	EH, 1u	20605·2	$a\,{}^1D_2{}^1$—72029·5
4793·800	A, 3u	20854·5	$a\,{}^3D_1{}^1$—72027·8
3982·4	Ha, 1u	25103·4	$a\,{}^3P_0$ —72024·6
3839·08	W	26040·5	$a\,{}^3P_1$ —72028·3
4227·93	H, 1 ; EH, 2	23645·6	$a\,{}^1F_3$ —73636·9
4043·50	H, 2 ; EH, 4 ; W	24724·1	$a\,{}^3D_3{}^1$—73636·6
3686·57	H, 2 ; EH, 3 ; W	27117·8	$a\,{}^3F_3$ —73636·9
2566·46	H, 1u ; EH, 1	38952·5	$a\,{}^1F_3$ —88943·8
2538·78	H, 1u ; EH, 1 ; W	39377·2	$a\,{}^3D_2{}^1$—88942·1
2437·95	W	41005·6	$a\,{}^3F_2$ —88944·9
2327·343	H, 1	42954·3	$a\,{}^3P_1$ —88942·1
2430·49	H, 1u ; EH, 1 ; E, 1u	41131·5	$a\,{}^1D_2{}^1$—92555·8
2415·93	W	41379·3	$a\,{}^3D_1{}^1$—92552·6
2348·80	H, 1u ; EH, 2 ; E, 3u ; W	42561·9	$a\,{}^1F_3$ —92553·2
2240·63	M	44616·4	$a\,{}^3F_2$ —92555·7
2146·93	M	46563·6	$a\,{}^3P_1$ —92551·4
2079·94	M	48062·9	$a\,{}^3P_2$ —92552·8
2336·26	H, 2u ; EH, 2 ; E, 4 ; W	42790·3	$a\,{}^1P_1$ —94457·4
2323·10	H, 1u ; EH, 1	43032·7	$a\,{}^1D_2{}^1$—94457·0
2309·61	H, 1u ; EH, 1 ; E, 2	43284·0	$a\,{}^3D_1{}^1$—94457·3
2226·84	H, 1u ; EH, 2 ; E, 3u ; W	44892·6	$a\,{}^3D_2{}^1$—94457·5
2000·71	M	49966·0	$a\,{}^3P_2$ —94455·9
2291·083	H, 2u ; EH, 2 ; E, 4 ; W	43634·1	$a\,{}^1P_1$ —95301·2
2278·449	H, 1u ; E, 2 ; W	43875·9	$a\,{}^1D_2{}^1$—95300·2
*2265·46	H, 1u ; EH, 2 ; E, 2 ; W	44127·5	$a\,{}^3D_1{}^1$—95300·8
2066·31	H, 1u	48379·9	$a\,{}^3P_0$ —95301·1

* Used twice.

Table II—(continued).

Wave-length.	Authority and Intensity.	Wave-number.	Terms.
2231·04	H, 1u ; EH, 1 ; E, 2	44808·2	$a\,{}^1D_2{}^1$—96232·5
2218·58	H, 1u ; EH, 1u	45059·8	$a\,{}^3D_1{}^1$—96233·1
2070·06	M	48292·3	$a\,{}^3F_2$ —96231·6
2027·19	M	49313·4	$a\,{}^3P_0$ —96234·6
2215·16	H, 2u ; EH, 2 ; E, 3 ; W	45129·4	$a\,{}^1P_1$ —96796·5
2046·08	M	48858·2	$a\,{}^3F_2$ —96797·5
1968·20	M (vac.)	50807·8	$a\,{}^3P_1$ —96795·6
1911·87	M (vac.)	52304·8	$a\,{}^3P_2$ —96794·7

Possibly genuine are the following:—

Wave-length.	Authority and Intensity.	Wave-number.	Terms.
2463·93	H, 1u	40573·3	$a\,{}^1P_1$ —92240·4
2205·62	M	45324·5	$a\,{}^3P_0$ —92245·7
2161·356	H, 2u ; EH, 2 ; E, 1u ; W	46254·7	$a\,{}^3P_1$ —92242·5
2421·94	H, 1	41278·1	$a\,{}^3D_1{}^1$—92451·4
2195·75	H, 2u ; EH, 2 ; E, 3 ; W	45528·3	$a\,{}^3P_0$ —92449·5
2151·83	H, 1u ; EH, 1 ; E, 2	46457·6	$a\,{}^3P_1$ —92445·4
2224·75	H, 1u ; EH, 2 ; E, 2 ; W	44937·1	$a\,{}^1P_1$ —96604·2
2212·85	H, 1R ; EH, 1 ; E, 1 ; W	45176·5	$a\,{}^1D_2{}^1$—96600·8
2200·57	H, 1u	45428·5	$a\,{}^3D_1{}^1$—96601·8
2359·00	H, 1u	42377·9	$b\,{}^3D_1$ —130531·2
*2265·46	H, 1u ; EH, 2 ; E, 2 ; W	44127·5	$b\,{}^3D_2$ —130532·1
2249·13	H, 2u ; EH, 2 ; E, 3u ; W	44447·8	$b\,{}^3D_3$ —130531·5

* Used twice.

Summary.

(A) By fusing copper wires in a vacuum grating spectrograph, a "snap-shot" of the copper spectrum is obtained. The lines 1358·76, 1367·92 and 1472·38 Å.U. are allocated to $1\,{}^1S_0 - a\,{}^1P_1$, $1\,{}^1S_0 - a\,{}^3D_1{}^1$, and $1\,{}^1S_0 - a\,{}^1P_1$ respectively. This identifies $1\,{}^1S_0$—the lowest term of the copper spark spectrum—with the term-value —21929·4.

(B) Photographs of the spectrum of copper-gold are taken with a small quartz spectrograph, using the Pfund arc. The wave-lengths of the copper and gold lines are measured, assuming as standards the calculated values of three copper lines. Thus it is shown that the Shenstone scheme of copper spark terms is consistent with the gold arc term-scheme of McLennan and McLay.

(C) Some additional terms have been found in the copper spark spectrum, and their values are given.

The Existence of Intermetallic Compounds in the Vapour State. The Spectra of the Alkali Metals, and of their Alloys with each other.

By J. M. WALTER, B.Sc., and S. BARRATT, B.A., University College, London.

(Communicated by T. R. Merton, F.R.S.—Received March 22, 1928.)

[PLATE 1.]

It has been uncertain until recently, whether any of the many intermetallic compounds which have been established in liquid and solid alloys can persist into the vapour state, or whether they invariably dissociate into mere mixtures of the individual metal vapours. Volatile compounds have occasionally been suspected (*e.g.*, between magnesium and zinc) on account of the deposition of typical crystals of compounds during the distillation of an alloy of the two metals, but such evidence is so indirect that it carries little weight. The most promising way of approaching this question is by a study of the absorption spectra of mixtures of metallic vapours. The metals that are known to possess a measurable proportion of diatomic or polyatomic molecules in the vapour state invariably indicate this property by the appearance of bands in their absorption spectra. In fact, since it has been accepted that band spectra cannot originate from single atoms; the observation of such bands has become the most delicate test for the existence of association in the vapours.

For example, the usual methods of molecular weight determination have led to the belief that the alkali metal vapours are solely monatomic. This cannot, however, be true, as each of these metals possesses an extensive band spectrum, which appears at quite low vapour pressures. There can be no doubt, on the basis of this spectroscopic evidence, that an appreciable fraction of the molecules in the vapour of boiling sodium, etc., are di- (or poly-) atomic. This point will be discussed fully below.

It seems probable that the same procedure will provide a very general test for the presence of molecules of intermetallic compounds persisting into the vapour state. The first of these compounds to be detected in this way was that between sodium and potassium.* It possesses one typical band spectrum in the yellow-green, and another (recorded below) in the ultra-violet. It has since been shown, by the appearance of characteristic bands in the vapours, that magnesium forms volatile compounds with each of the alkali metals.* We have now extended the observations to include the vapours of all the possible pairs of alkali metals. With one exception (lithium-sodium) we have found that each pair of metals possesses a characteristic band absorption spectrum, so that the power of forming volatile compounds with each other must be a general one within this group. The details of these new spectra are given in a later section.

---

* 'Roy. Soc. Proc.,' A, vol. 105, p. 221 (1926).

In establishing the existence of an intermetallic compound by this method, the assumption is made that the spectrum of a mixture of two metals is purely additive unless a new type of molecule (*i.e.*, of a compound) is formed, which is peculiar to the mixture. This assumption is justified by all the evidence available. Further verification follows from the intensity variation of the new spectrum with the relative amounts of the two components of the vapour.

If the compound molecules are diatomic, then from mass law considerations the intensity of the bands should be the same from a mixture of a small amount of the component A with large excess of component B, as from a small amount of B with large excess of A. This method of verification has been applied successively to all the spectra described in this paper.

Independent confirmation of the conclusions reached by these methods can be obtained by an analysis of the fine structure of the bands. Thus H. G. Smith† examined the sodium, the potassium, and the sodium-potassium spectra under high dispersion, and deduced the moments of inertia of the corresponding molecules. The values so obtained were Na_2 $2 \cdot 3 \times 10^{-39}$; K_2 $18 \cdot 4 \times 10^{-39}$; NaK $6 \cdot 6 \times 10^{-39}$ gm./cm.^{2}.‡ Other factors besides the masses of the constituent atoms determine these moments of inertia, so that it could not have been predicted that the value for NaK would be intermediate between those for Na_2 and K_2. That it should prove to be so is, however, good confirmation of the suggested origin of the spectrum. The values are of the magnitude that would be expected for the molecules Na_2, K_2, NaK from crystallographic data.

The molecules of the alkali metals, and of the intermetallic compounds discussed in this paper, are therefore probably diatomic. This conclusion will be used in referring to them in the succeeding sections.

It is desirable, in work of this kind, to use several specimens of each metal, from independent sources, to ensure that no impurities can be responsible for the phenomena. The alkali metals in the present experiments were each prepared by several different methods from independent samples. We have actually found no difficulties with impurities in dealing with the alkali metals, but in further experiments we have proved that it is especially easy to draw false conclusions, unless this precaution is taken, in assigning the molecular spectra of the elements of the second group of the Periodic Table.

---

* 'Roy. Soc. Proc.,' A, vol. 109, p. 194 (1925).

† 'Roy. Soc. Proc.,' A, vol. 106, p. 400 (1925).

‡ See also Pringsheim and Rosen, 'Z. f. Physik,' vol. 43, p. 519 (1927).

The Proportion of Molecules present in Alkali Metal Vapours.

While the existence of molecules in alkali metal vapours is well established, the proportion of them present in the vapour is not yet known with any degree of accuracy. The early molecular weight determinations on the metals merely proved that the proportion could not be very large ; a small percentage of double molecules would have escaped detection. There have been two recent spectroscopic estimates of the concentration of the K_2 molecules in potassium vapour. Carelli and Pringsheim* determined the intensity of the potassium band spectrum, excited as fluorescence in the vapour by white light, and from their results calculated that the ratio $[K_2]/[K] = 2 \times 10^{-2}$ at 727° C. Ditchburn† assumed that the absorption of the diffuse band in the yellow-green is associated with the splitting up of the molecule into two atoms, one normal and the other in an excited state, and was able in this way to calculate the heat of formation of the molecule. The $[K_2]/[K]$ ratio at 727° C., using this value of the heat of formation, should be approximately $0 \cdot 9 \times 10^{-2}$.

In an attempt to measure the same ratio directly, a redetermination of the vapour density of potassium by the Victor Meyer method has been made in this laboratory by Mr. E. K. Mitchell. The experiments had unfortunately to be discontinued before they had been taken to their limit of accuracy, but they were sufficiently complete to warrant a brief description here. The following account has been included with Mr. Mitchell's kind permission :—

The Victor Meyer method was applied to the alkali metals by Dewar and Scott‡ but with indecisive results, owing to the difficulties of finding a suitable containing vessel and an inert gas. Preliminary experiments showed us that potassium vapour does not react (or only negligibly slowly) with an iron containing vessel at 1000° C., while argon, unknown at the time of the earlier work, solved the second difficulty.

The Victor Meyer apparatus itself consisted of a cylindrical bulb, made of mild steel, and of about 200 c.c. capacity. Into this was welded a tube 1 yard long and ½ inch diameter. Near the top, this tube carried a branch into which was fixed the apparatus for holding and releasing a steel container for the potassium. Another side tube led to a gas burette in which the displaced gas was measured. The apparatus was heated in an electric furnace, which was so heavily constructed that the temperature was easily kept constant during the progress of an experiment. The determinations were made at 950° C.

---

* 'Z. f. Physik,' vol. 44, p. 643 (1927).

† 'Roy. Soc. Proc.,' A, vol. 117, p. 486 (1928).

‡ 'Roy. Soc. Proc.,' A, vol. 29, p. 490 (1879).

The potassium was weighed in a small steel bottle, 1·5 cm. long and 0·7 cm. diameter, with a ground-in steel stopper. To introduce the metal into the furnace tube, this stopper was attached by a dovetailed groove to the internal piece of a brass ground joint (as shown in the diagram, fig. 1) and the whole inserted into the branch tube of the main apparatus. On turning the ground joint, two lugs on the end of the bottle engaged with spiral projections inside the main tube (shown in cross section only), and the bottle itself, being forced off, fell into the furnace bulb leaving the stopper behind. This method ensured the maximum rate of volatilisation of the metal after dropping. After an experiment, the weighing bottle was recovered from the furnace bulb by "fishing" for it with an electromagnet. The accurate weighing of the potassium which was to be volatilised was the greatest difficulty we encountered. Oxide formation, if permitted, would have given spuriously high values of the vapour density. We finally devised a fairly simple method of transferring the pure metal in safety to the weighing bottle. The principle of this will be obvious from fig. 2. A considerable quantity of potassium was distilled *in vacuo*

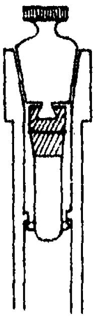


FIG. 1.

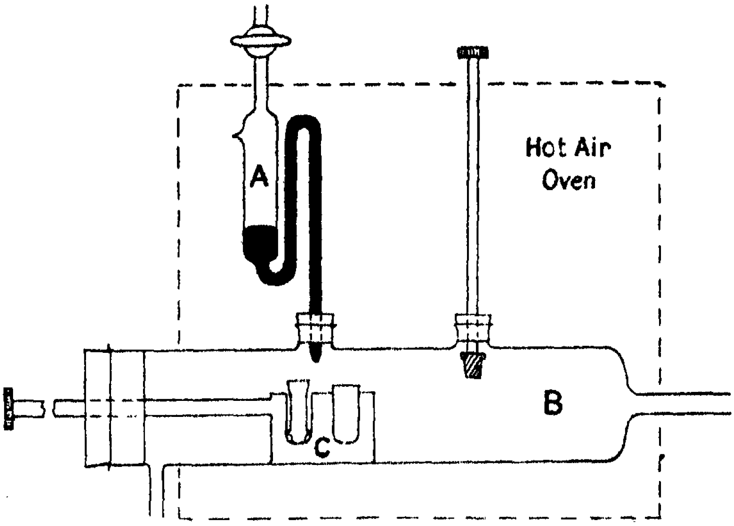


FIG. 2.

through a series of bulbs and ultimately into the tube A, which was then filled with argon. A had a siphon exit tube, which was sealed during the distillation and was opened afterwards in an inert atmosphere. The open nozzle of this siphon was fixed by a rubber stopper into the tube B. Another side tube in B accommodated an apparatus for replacing the stopper in the weighing bottle after the charging. The weighing bottle and another dummy container were fixed into the trolley C. A slow current of hydrogen (passed over heated copper, and through long P_2O_5 drying tubes) flowed through B continuously. The whole apparatus was heated in an air oven to a temperature above the melting point of potassium when the metal could be expelled from the siphon by increasing the gas pressure in A. The first drops falling were rejected into the dummy container and the weighing bottle was then charged. The trolley was next pushed into a suitable position for the insertion of the stopper. The drop of potassium hanging on the end of the siphon tube remained untarnished for many minutes when the hydrogen stream was really dry.

Silica, etc., had been removed from the interior of the furnace bulb by fusing caustic potash inside it. Dry hydrogen was passed through the heated bulb for some hours before each experiment, and was replaced by argon immediately before the actual determination. The argon was taken from a cylinder and purified over heated sodium and P_2O_5 before entering the tube. In short, stringent precautions were taken to avoid the contact of oxygen, or of oxides, with the metal both before and *after* volatilisation. Our final value for the molecular weight of potassium, so determined, was 41·6 ($T = 935°$ C.). The work had most unfortunately to be abandoned before it had run its normal course of confirmatory experiments, and we can only deduce with safety that the ratio $[K_2]/[K]$ at 935° C. and atmospheric pressure does not exceed 5×10^{-2}. It may be mentioned that preliminary trials of the apparatus with mercury vapour gave results within ½ per cent. of the value for the monatomic vapour. This direct determination of the molecular ratio agrees well with the indirect estimates quoted above. The amount of the molecular form is surprisingly small in view of the great intensity of the banded absorption at quite low pressures.

Though potassium is the only metal for which the molecular ratio has been determined, this quantity is probably of the same order for the other metals of the group, if one may judge from the relative intensities of their band spectra, seen through vapours of equal density. The conclusion is therefore that the alkali metals at their boiling points contain not more than 5 per cent. and probably only about 2 per cent. of molecules of the type M_2.

No determinations of the equilibrium amounts of the "mixed molecules" of the type NaK have yet appeared, but we believe that in an equimolecular mixture of sodium and potassium the concentrations $[Na_2]$, $[K_2]$, and $[NaK]$, are all of the same order. This opinion is based on a series of experiments upon the minimum temperature at which the bands could be detected photographically in the absorption of a column of saturated vapour of definite length. A hard glass tube about 8 inches long was filled with an equimolecular mixture of sodium and potassium, and gradually heated in an electric furnace. A series of photographs was taken as the temperature rose, and the first appearance of the various band systems was picked out. Potassium is the more volatile metal, and the red K_2 bands were the first to appear, then the NaK bands, and finally the green Na_2 bands. Bearing in mind the mass action expression $[NaK]/[Na] \times [K] =$ constant, this observation suggests at once that NaK molecules are formed about as readily as either Na_2 or K_2. The spectra recorded below are therefore due, it would seem, to molecules of intermetallic compounds which form approximately 1 per cent. of the total vapour present in the absorbing column.

The Band Spectra of the Alkali Metals.

Our experimental procedure in studying the band spectra of the alkali metals, and of their alloys, has already been fully described in another connection.[*] The absorbing column of vapour was about a foot long, and was produced in a steel tube electrically heated. The ends of the tube were covered with quartz windows, and were water-cooled. Excepting during the exposures, the windows were protected from the vapours by internal glass shutters. Without this protection they were apt to cloud over by condensation of the vapours before the end of the observations. A 500-c.p. Pointolite lamp formed the continuous background in the visible (and on occasion down to λ 3100), and the crater of a carbon arc served for the ultra-violet. There is little difficulty in reaching λ 2200 by this method. The spectrographs used were two Hilger instruments, one of the constant deviation type and the other an "E.3" model. An atmosphere of argon was always present in the furnace tube to slow up the rate of distillation of the metals.

The lithium, sodium and potassium used in the experiments were ordinary laboratory samples. Rubidium and cæsium are difficult to handle in the metallic form, and were prepared *in situ* in the furnace tube by the action of some other metal on the chlorides. In studying the spectrum of a mixture such as that of lithium and potassium, some little difficulty was found in producing the mixed vapours in the tube. The boiling point of lithium is so high (1400°) in comparison with that of potassium that the latter distils away from the heated portion of the tube before the lithium has begun to volatilise. The solution found was to introduce the metals into the tube in separate iron containers. The lithium was placed at once in the hot part of the tube, while the potassium rested over one of the water-coolers. When the lithium was volatilising freely, the potassium was drawn into the heated section by a steel wire attached to the container. This wire was manipulated from outside through a mercury seal.

---

[*] O. H. Walters and S. Barratt, 'Roy. Soc. Proc.,' A, vol. 118, p. 120 (1928).

As a preliminary to the recognition of the spectra of the intermetallic compounds, we made a thorough examination of the band spectra of the individual metals. We found that these latter spectra have been very incompletely described, particularly those of rubidium and cæsium. In this section it is proposed to deal in turn with the spectra of the metals taken separately and then with those of the compounds formed between them. The new bands that we have ascribed to rubidium and to cæsium were unaffected by the method employed for the preparation of the metals. They were all observed equally well when the metal vapour was prepared by reduction of the chloride with calcium, magnesium, or lithium. Their intensities were unchanged by several successive distillations of the metals *in vacuo*, and as substitution of a hydrogen atmosphere for the argon in the tube had no influence on the relative intensities of the various bands, it may be assumed that none of them are hydride bands. Neither would their structure suggest such an origin. All the spectra have been examined down to about λ 2300. Unless otherwise stated, the bands are degraded to the long wave-length side. The bands marked with an asterisk are those which are difficult to measure satisfactorily.

Lithium.

Bevan* attempted to obtain the absorption spectrum of lithium, but was unsuccessful, owing to the rapid action of the vapour on his furnace tube. Narayan and Gunnaiya† succeeded in photographing an extensive band region in the green, and measured some 33 heads. There are also some weak bands in the ultra-violet, but these we found unsuitable for measurement under a low dispersion.

---

* 'Roy. Soc. Proc.,' A, vol. 84, p. 213 (1910).

† 'Roy. Soc. Proc.,' A, vol. 106, p. 51 (1924).

Sodium.

This band spectrum has attracted many investigations. R. W. Wood, in a well-known series of papers, has described this spectrum very fully down to λ 4100, but below this wave-length we can find no measurements on the band absorption, though both Wood and Bevan mention a band region in the neighbourhood of the second member of the principal series λ 3303, while Wood states that he has observed bands accompanying each member of the principal series down to λ 2594. We have repeated these observations and have measured the heads in each region. All the bands are degraded to the long wave-length side. The system near λ 3303 contains three groups of bands with characteristic

Measurements on the Sodium Band Heads.

λ_{air}	$\nu_{vac.}$	Int.	$\lambda_{air.}$	$\nu_{vac.}$	Int.
3526	28353	1	3188	31359	1
3505	28523	1	3185	31388	1
3486	28678	4	3181	31428	1
3468	28827	8			
3450	28977	10	3073	32532	1
3432	29129	10	3059	32681	1
3414	29283	8	3045	32831	1
3400	29403	8	3031*	32983	1
3383	29551	8	3020	33103	1
3370	29665	6	3006	33257	1
3357	29780	6	2992	33413	1
3352	29824	2	2982	33525	1
3345	29887	1	2972	33638	1
3339	29941	3	2968	33683	1
3326	30058	3	2962	33740	1
3315	30157	1	2958	33797	1
	(λ 3303 absorption)		2954	33843	0
3280	30479	4	2949	33900	1
3274	30535	2	2940*	34004	1
3268	30591	4	2931	34108	1
3263	30638	4	2922	34213	1
3257	30694	4	2913	34319	1
3252	30741	4	2905	34413	1
3247	30789	2	2896	34520	1
3242	30836	1			
3237	30884	1	2839*	35213	0
3235	30903	1	2830*	35325	0
3231	30941	1	2819*	35463	0
3227	30979	1	2811*	35564	0
3223	31018	1		(Emission line.)	
3221*	31037	1	2788*	35857	0
3213	31115	1	2781	35948	0
3210	31144	1	2773	36051	0
3207	31173	2	2758	36247	0
3204	31202	2			
3200	31241	2	2667	37484	1
3197	31270	2	2580	38613	1
3194	31300	2	2470	40474	1
3191	31329	2			

wave-length intervals of about 20 A., 6 A., and 3 A. respectively. The system near λ 2853 is weaker. It consists of two groups with spacings of about 14 A. and 10 A. The bands near λ 2680 are still weaker, with a spacing of 10 A. In addition to these groups we have observed three isolated bands at λλ 2667, 2589, and 2470. We have also confirmed Wood's observation that with dense vapour a continuous absorption extends from λ 3303 to the far ultra-violet.

Potassium.

In 1908 Wood and Carter[†] found a band absorption in this vapour between λ 6667 and λ 6346. This spectrum was described independently by Bevan.[‡] Datta[§] photographed a banded absorption between λ 4330 and λ 4610. This was also observed by McLennan and Ainslie,[||] but no measurements appear to have been put on record. McLennan and Ainslie have also observed and measured a further absorption region between λ 8612 and λ 8213.

We have measured the main heads in the band about λ 4230, and even at high vapour densities no further bands could be detected, with the exception of a diffuse band with a maximum at λ 5770. This is the beginning of a weak continuous absorption stretching into the blue.[¶] The nature of this band will be discussed in a later section.

Measurements on the Potassium Band Heads.

λ_{air}.	ν_{vac}.	Int.	λ_{air}.	ν_{vac}.	Int.
4612	21677	0	4365	22903	1
4597	21747	0	4360	22929	3
4582	21818	1	4348	22993	5
4565	21900	1	4337	23051	2
4546	21991	0	4331	23083	1
4521	22113	0	4325	23115	0
4498*	22226	1	4320	23142	6
4479*	22320	1	4308	23206	8
4460	22415	0	4297	23266	7
4425	22593	2	4287	23320	5
4406	22690	3	4276	23380	2
4389	22778	1	4266	23435	1
4377	22840	4	4250	23523	0
4371	22872	0	4243	23562	0

[†] 'Phys. Rev.,' vol. 27, p. 107 (1908).

[‡] 'Roy. Soc. Proc.,' A, vol. 84, p. 213 (1910).

[§] 'Roy. Soc. Proc.,' A, vol. 101, p. 539 (1922).

[||] 'Roy. Soc. Proc.,' A, vol. 103, p. 304 (1923).

[¶] 'Roy. Soc. Proc.,' A, vol. 105, p. 221 (1924).

Rubidium.

Cater† and Bevan‡ described a band region between λ 6700 and λ 7100. McLennan and Ainslie (*loc. cit.*) also observed this absorption and measured four more bands. In addition, we have now observed an absorption in the orange-red, consisting of three strong heads degraded to the long wave-length side. There are also five diffuse and structureless bands which we attribute to this spectrum, as they satisfy the criteria mentioned above. The band at λ 4746 is much stronger than the others, which only appear with very dense vapour. The new measurements are as follows :—

λ_{air}.	ν_{vac}.	Int.	
6054	16513	5	
6033	16571	5	
5992	16684	5	
5081	19676	0	Maximum of diffuse band.
4746	21065	2	,, ,,
4703	21257	0	,, ,,
4288*	23314	0	,, ,,
4280	23358	0	,, ,,
4103	24366	00	,, ,,

Cæsium.

McLennan and Ainslie (*loc. cit.*) observed five bands between λλ 7585 and 7778. Liveing and Dewar§ refer to an absorption in the yellow-green, which they ascribe either to cæsium or to the mixture sodium-cæsium. By using carefully purified cæsium chloride and calcium that was almost sodium-free, we have been able to show that this absorption is not due to cæsium vapour. It is the spectrum of the sodium-cæsium molecule, and is described in the appropriate section below.

In the vapour of pure cæsium, we have found, in addition, four narrow bands near λ 7000, and an extensive band system in the near red, coming to a head at λ 6265. The intensities of the bands in this system increase smoothly up to the head (see Plate 1, fig. 1). Finally two structureless bands are also listed. They are of approximately equal intensities.

---

† 'Phys. Zeit.,' vol. 11, p. 632 (1910).

‡ 'Roy. Soc. Proc.,' A, vol. 85, p. 61 (1911).

§ 'Collected Papers,' p. 12.

Measurements on the Cæsium Band Heads.

$\lambda_{air.}$	$\nu_{vac.}$	$\lambda_{air.}$	$\nu_{vac.}$
7187	13910	6510·0	15356·7
7126	14029	6489·8	15404·5
7071	14138	6468·0	15456·5
7051	14178	6449·3	15501·3
		6428·5	15551·4
6890·7	14508·3	6408·5	15600·0
6843·8	14607·7	6389·0	15647·6
6810·3	14679·6	6371·0	15691·8
6767·3	14772·9	6354	15734
6725·5	14864·7	6335	15781
6682·5	14964·6	6321	15816
6659·3	15012·5	6301	15866
6636·3	15064·5	6282	15914
6615·8	15111·2	6265	15957
6593·5	15162·3		
6572·3	15211·3	5224	19139
6552·0	15258·3	4718	21190
6531·5	15306·2		

Lithium-Sodium.

We have not been successful in photographing any bands that can be attributed to this molecule. The band systems of lithium and of sodium are themselves extensive, so that it is not unlikely that such compound bands do exist but are invariably masked by the spectra of the Li_2 and the Na_2 molecules.

Lithium-Potassium.

A new system of bands appeared in the mixture, on the short wave-length side of the D lines. They are more widely spaced than the NaK bands which are in the same region (see Plate 1, fig. 2).

Measurements.

$\lambda_{air.}$	$\nu_{vac.}$	Int.
5838·7	17122·4	3
5769·3	17328·3	3
5724·0	17465·5	2
5700·0	17539·0	2
5658·0	17669·2	3
5620·0	17788·7	3
5583·0	17906·6	1
5519·5	18112·6	0

Lithium-Rubidium.

There is also a new band spectrum characteristic of this mixture. Owing to traces of potassium in the rubidium the strongest members of the LiK spectrum were always superimposed on these bands, but they were easily identified.

Measurements.

λ_{air}.	ν_{vac}.	Int.
5815·5	17192·1	1
5778·0	17302·2	2
5743·0	17407·7	2
5712·5	17500·6	1
5684·0	17588·4	0

Lithium-Cæsium.

The LiCs spectrum consists of a band system in the red.

Measurements.

λ_{air}.	ν_{vac}.	Int.
6255	15983	4
6217	16080	3
6180	16177	4
6146	16266	4
6116	16346	3
6084	16432	1
6055	16511	0
6029	16582	0

Sodium-Potassium.

The bands of the NaK spectrum in the yellow-green have already been described.* It is interesting to note that they were accidentally observed by Roscoe and Schuster,† who believed that they were part of the potassium spectrum. The matter is explained by an analysis they give of the potassium used in the experiments, for 0·8 per cent. of sodium was present in it. These bands have since been examined, as stated above, by H. G. Smith. Pringsheim‡ has observed them in the fluorescence spectrum of sodium-potassium mixtures. Newman§ has also described bands in much the same spectral region, which he found in the arc spectrum of this mixture. These bands are not the same, however, as those under discussion.

---

* Barratt, *loc. cit.* (1924).

† 'Roy. Soc. Proc.,' A, vol. 22, p. 362 (1874).

‡ 'Z. f. Phys.,' vol. 38, p. 161 (1926).

§ 'Phil. Mag.,' vol. 48, p. 159 (1924).

We have now extended the observations to the ultra-violet, and a second extensive band system of this molecule has been found; the wave-lengths are given below.

Second Group of Sodium-Potassium Band Heads.

$\lambda_{air.}$	$\nu_{vac.}$	Int.	$\lambda_{air.}$	$\nu_{vac.}$	Int.
4014	24906	1	3768	26532	2
4008	24943	1	3762*	26574	1
4001	24987	1	3756	26617	1
3998	25005	1	3751	26652	0
3989	25062	2	3744	26702	1
3976	25144	4	3739	26738	0
3969	25188	4	3735	26766	1
3958	25258	8	3728	26816	1
3944	25348	4	3717	26896	4
3940	25374	4	3708	26961	4
3930	25438	2	3700*	27019	0
3923	25484	8	3697	27041	1
3905	25601	10	3689	27100	1
3899	25640	4	3685	27129	1
3892	25686	4	3680	27166	0
3886	25726	10	3673	27218	0
3880	25766	2	3669	27248	2
3874	25806	8	3664	27285	0
3868	25846	4	3660	27315	1
3861	25893	2	2654*	27360	1
3856	25926	4	3650	27389	2
3850	25967	2	3641	27457	4
3844	26007	2	3632	27525	4
3839*	26041	1	3622	27601	6
3832*	26089	1	3612	27678	4
3826	26130	2	3607	27716	1
3821	26164	2	3603	27747	2
3814*	26212	1	3595	27809	2
3808	26253	1	3586	27878	4
3802	26295	1	3577	27948	4
3797	26329	1	3569	28011	4
3793	26357	1	3561	28074	2
3785	26413	1	3553	28137	2
3779	26455	2	3545	28201	1
3774*	26490	1	3538	28257	1

Sodium-Rubidium.

The NaRb spectrum lies in the green. We have measured the following five heads:—

$\lambda_{air.}$	$\nu_{vac.}$	Int.
5432	18404	4
5397	18524	4
5368	18624	4
5339	18725	4
5331	18753	4

Sodium-Cæsium.

The NaCs spectrum is rather an extensive one, in the yellow-green (Plate 1, fig. 3). This is the spectrum seen by Liveing and Dewar, as has been mentioned under "Cæsium."

Measurements on the Sodium-Cæsium Band Heads.

λ_{air}.	ν_{vac}.	Int.	λ_{air}.	ν_{vac}.	Int.
5732	17441	1	5512	18137	7
5710	17508	1	5492	18203	10
5702	17533	1	5483	18233	3
5672	17626	5	5473	18266	3
5642	17719	7	5463	18300	10
5631	17754	10	5444	18364	10
5610	17820	1	5425	18428	10
5602	17846	10	5406	18493	2
5591	17881	3	5388	18555	0
5571	17945	10	5367	18627	5
5562	17974	3	5353	18676	5
5542	18039	10	5338	18728	5
5521	18108	5			

In the ultra-violet Bevan[*] has ascribed a band spectrum in the neighbourhood of λ 3303 to this mixture. This is really part of the Na_2 spectrum. We have found, however, that there is a second spectrum of the NaCs molecule between λ 4315 and λ 3930 (see Plate 1, fig. 6).

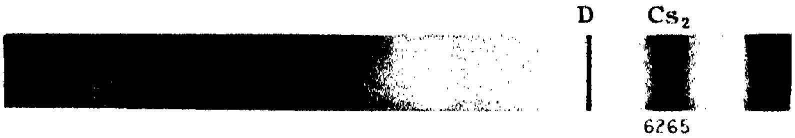
Second Group of Sodium-Cæsium Heads.

λ_{air}.	ν_{vac}.	Int.	λ_{air}.	ν_{vac}.	Int.
			4098	24395	10
4315	*23168*	*0*	4087	24461	6
4296	23271	0	4079	24509	6
4275	23385	0	4070	24563	6
4258	23479	1	4061	24618	4
4242	23567	2	4053	24666	2
4224	23668	4	4034	24782	4
4208	23758	4	4026	24832	4
4197	23820	1	4015	24900	4
4190	23860	4	4006	24956	6
4180	23917	6	3998	25005	6
4174	23951	2	3989	25062	6
4163	24014	6	3980	25119	6
4152	24078	6	3974	25156	4
4147	24107	4	3966	25207	1
4136	24171	10	3960	25245	1
4129	24212	4	3952	25297	1
4125	24236	6	3945	25341	0
4119	24270	6	3938	25386	0
4108	24336	6	3931	25432	0
4102	24371	6			

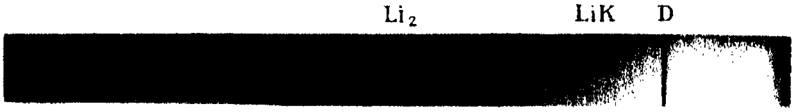
---

[*] 'Roy. Soc. Proc.,' A, vol. 83, p. 421 (1910).

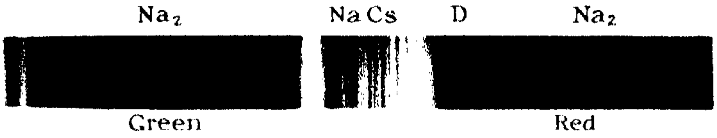
Walter and Barratt. *Roy. Soc. Proc., A, vol.* 119, *Pl.* 1.



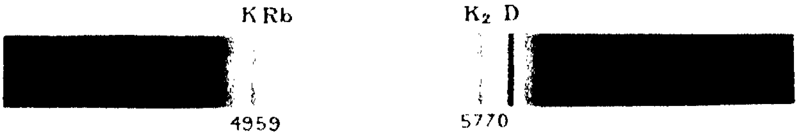
1



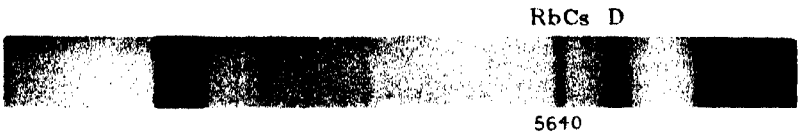
2



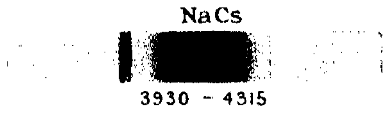
3



4



5



6

Potassium-Rubidium.

Bevan* believed that two bands in the green were characteristic of this mixture. We have found, however, that the only spectrum of this molecule is a structureless band with a maximum at λ 4959 (see Plate 1, fig. 4).

Potassium-Cæsium.

No bands with fine structure appear to be characteristic of this mixture. As with KRb, there is a new diffuse band. Its maximum is at λ 5387.

Rubidium-Cæsium.

This spectrum, which has been recorded by Dunoyer† also consists of a diffuse band, with a maximum at λ 5640 (see Plate 1, fig. 5).

Discussion.

The new bands described in the preceding sections add very materially to the number of the known alkali metal spectra. It seems probable that each molecule of the type MM′, where M and M′ are alkali metal atoms, possesses at least one system of vibrational bands. A few of these spectra we have not observed, but the probable overlapping of adjacent band systems, such as the extensive ones of Li_2 and Na_2, provides a very plausible explanation of these few failures. Quite a long series of analogous molecules is therefore available for comparison, and with further study of these spectra it should be possible to elucidate the structure of this type of molecule very considerably. Such loosely bound molecules are likely to be unusually simple. Each atom has a solitary valency electron, and the molecules are homopolar. There has been no detailed discussion of molecules so constituted. Franck‡ suggested that the atoms are here bound together by " Van der Waals " forces, and that there is no sharing of electronic orbits. On this view, the orbits which exist in the atoms should persist in the molecules, though in a disturbed form. As the production of the visible and ultra-violet bands of the molecules involves the excitation of one of the valency electrons, the bands should always lie in the neighbourhood of a principal series line, as these lines form the atomic absorption. This inference is borne out by the position of the sodium bands, for bands accompany at least the first four members of the principal series. The potassium molecule has band systems fairly near the first two principal lines, and rubidium and cæsium both have band systems in the red. The lithium bands, however, stand in no obvious relation to the atomic lines. It is interesting to note how far the new compound bands fit in with the positions predicted for them in this way. Bands would be expected near the series lines of both the constituent atoms. This expectation is realised by the spectra of the molecules NaK and NaCs. Both these compounds have one band system lying very close to the D lines, and another system in the blue, in the neighbourhood of the second member of the principal series of potassium and of cæsium respectively. None of the other molecules give more than one band system, so far as our observations can show. The molecules containing a lithium atom, like the lithium molecule itself, appear to be definite exceptions to the rule, as most of the spectra lie in the yellow, away from any of the atomic lines. If there is otherwise general obedience to this rule for the position of the bands, it will help to account for the missing spectra, as the chances of overlapping of the different systems will be much increased.

---

* *Loc. cit.* (1910).

† 'Le Radium,' vol. 9, p. 218 (1912).

‡ 'Trans. Faraday Soc.,' vol. 21, p. 536 (1926).

Following up the same conception of these molecules, each band system should possess a "Konvergenzstelle" on the short wave-length side of a series line. The frequency at this convergence point should give, by the usual quantum relationship, the energy required for the dissociation of the molecule into one normal and one excited atom. The frequency difference between the series line and the convergence point would therefore determine the heat of formation of the molecule. Unfortunately, none of the well-developed band systems of these molecules seem to be suited for an application of this theory. $\Delta\nu$ is usually almost constant throughout a system (or takes one or two definite values), so that over the range of frequencies that can be observed, the molecules appear to be acting very nearly as harmonic oscillators. It may be noted in passing how frequently $\Delta\nu$ takes the values 100, 50, or 30, in these spectra, irrespective of the constituent atoms of the molecule and of the position of the bands. Without exception the bands are degraded to the red, so that, with increasing electronic excitation, the moment of inertia of the molecules—and the distance between the atomic centres—must decrease.

We have found that most of the alkali metal molecules possess a distinct type of absorption spectrum, in addition to these band systems. This second type consists of an isolated and diffuse band, which shows no signs whatever of fine structure under the dispersion we have been able to apply, and which is probably a true continuous spectrum. Usually there is only one of these bands in the spectrum of each molecule, but cæsium possesses two of approximately equal intensity, $\lambda\lambda$ 5224 and 4718. The bands are not very broad—perhaps 30 A. or so on our plates—but with some of them we have been able to detect a faint continuous absorption running towards the ultra-violet. The K_2 and the KRb bands of this type are shown in Plate 1, fig. 4, and the RbCs band in fig. 5, but the weak continuous absorptions connected with them have been lost in reproduction. The bands are not strongly developed, except in rather dense vapours, but we have not experimented seriously upon their variation with the pressure of the metallic vapour. They are not, however, visibly affected by the presence, or absence, of an atmosphere of argon or hydrogen.

It is a tempting hypothesis that these bands are analogous to the well-known continuous absorption bands of the halogen molecules. The explanation of the existence of these was first given by Franck (*loc. cit.*). They correspond to the dissociation of the molecule into two atoms, one excited and the other normal, with more or less kinetic energy. If this view of their origin is accepted, the long wave-length edges of the diffuse bands of the alkali metals should coincide with the "Konvergenzstelle" of the vibrational bands. This confirmation of their origin, as we have seen above, unfortunately cannot be obtained, as the bands show no sign of convergence where measurements are possible. We may, however, calculate the heat of formation of the molecules from the position of the diffuse bands, and compare the values with estimates from other data. Even this test can only be applied to the potassium spectrum, as the heats of formation of the other molecules are as yet unknown. Ditchburn (*loc. cit.*) has made photometric measurements on the continuous absorption of potassium in the visible, and, using Franck's interpretation, has calculated the energy of formation of the K_2 molecule to be 0·51 volt. Carelli and Pringsheim (*loc. cit.*) have deduced values of 0·51–0·63 volt, from measurements on the intensity of fluorescence in the vapour. Our own value from the position of the diffuse band of potassium is almost identical with that of Ditchburn (see the table below), and it is very probable that the long wave-length end of his continuous absorption coincides with the band in question. The agreement of the two independent values is quite satisfactory for this spectrum, especially in view of the difficulty of the fluorescence measurements.

Another test of the hypothesis would be afforded by the appearance of a similar "diffuse" band at the same frequency interval from the second member of the principal series. The predicted position of this band is approximately λ 3440 for the potassium spectrum. There is undoubtedly a diffuse band at λ 3430 in dense potassium vapour, which may be the absorption for which we were searching, but we do not think the identification at all certain, owing to the proximity of the principal pair at λ 3447. These lines, as is well known, develop satellites in dense vapour, and the band may merely be one of these.

We have tested similar predictions on the other spectra, where possible, but with negative or indecisive results. The intensity of such bands should, however, be low in any case, and their total absence would not be surprising.

We have tabulated below the approximate heats of formation of some of the alkali molecules, based on the assumption of this origin for the bands.

The sodium band listed is the one which develops in the green in dense vapours, and is apparently structureless. We have used Wood's measurement of this band. We have given two values for the cæsium molecule as there are two diffuse bands in its spectrum, of almost equal intensity, and we have no immediate means of choosing between them.

Molecule.	Band.	Heat of formation.	
		Volts.	Calories.
	λ		
Na_2	5530	0·14	3,200
K_2	5770	0·53	12,000
Rb_2	4746	1·02	23,000
Cs_2	5224	0·91	21,000
	4718	1·17	27,000
KRb	4959	0·89	20,000
KCs	5387	0·78	18,000
RbCs	5640	0·71	16,000

Possibly examination of both under high dispersion may settle the point, but the existence of two such bands is certainly a difficulty for the hypothesis at present. The value, too, for sodium is unexpectedly low, in view of the great intensity of its band spectrum. It is evident, from the occurrence of band systems at higher frequencies than these diffuse bands, and on other grounds, that the molecules can take up energy well in excess of their requirements for dissociation and still remain intact. To quote Franck's view again, it is probably only when the vibrational energy exceeds the limiting value that dissociation is certain. In fact, as we have seen, the electronic excitation of one of the atoms actually causes the nearer approach of the atomic centres. The molecules, after absorption of visible or ultra-violet light corresponding to a band frequency, may dissociate, as Wood, and also Ditchburn, have shown, by observing the emission of the D lines in the fluorescence accompanying such band absorption. They may also fall back to the normal state, for it is possible to observe the bands in emission. Wood mentions this fact and we have verified it for the Na_2, the K_2, and the NaK spectra, by applying electrical excitation.

Which of the two mechanisms predominates, presumably depends on the collision frequency.

Similar " diffuse " bands have been found in the spectra of the magnesium alkali molecules,* and we have also observed them during an examination, as yet incomplete, of the alloys of mercury and cadmium with the alkali metals. It is not unlikely, therefore, that they will prove to be a general feature of the spectra of metallic molecules. If the above hypothesis of their origin is finally confirmed, it will be possible to estimate the heats of formation of many such compounds from the position of their bands. As an example, we may take the sodium-magnesium molecule. The spectrum of this compound consists of a structureless band with a sharp boundary on the long wave-length side at λ 5291. Assuming that the molecule breaks up on absorbing light of this wave-length into a normal magnesium atom and an excited sodium atom, the heat of formation is approximately 0·24 volt, or 5500 cals. We hope to be able to give data for other pairs of metals in a subsequent communication.

We wish to thank Mr. E. K. Mitchell, B.Sc., very heartily for his permission to include an account of his experiments on the vapour density of potassium.

Summary.

(1) The alkali metals, Li, Na, K, Rb, and Cs form a complete series of binary molecules with each other in the vapour state. At the boiling point of the metals, these molecules constitute 1–2 per cent. of the total vapour.

(2) The spectra of these compounds are described. The band systems lie, in general, near the lines of the principal series of the constituent atoms. The molecules containing lithium are exceptions to this rule.

(3) In addition to these bands, narrow regions of continuous absorptions are typical of the absorption spectra of the molecules. It is suggested that these regions correspond to the continuous absorptions of the halogen molecules, and the heats of formation of the alkali metal molecules have been worked out on this basis.

(4) The vapour density of potassium has been redetermined by the Victor Meyer method. The proportion of the molecular form cannot exceed 5 per cent.

---

* Barratt, *loc. cit.* (1925).

The Theory of Rayleigh's Principle as applied to Continuous Systems.

By G. Temple, Ph.D., Imperial College.

(Communicated by S. Chapman, F.R.S.—Received March 23, 1928.)

1. *Rayleigh's Principle.*

The object of this paper is to examine and extend a method frequently applied by Lord Rayleigh to the calculation of the frequency of the gravest mode of a vibrating system. In this method no attempt is made to obtain an accurate solution of the differential equations of vibration in any normal mode, but any fairly simple function satisfying the boundary conditions is adopted as an approximate solution. This solution will involve the time only through a harmonic factor such as $\sin 2\pi\nu t$, so that the mean value of the kinetic and potential energies of the system may be calculated—always on the supposition that frictionless constraints are applied so that the displacements of the system follow the law of the approximate solution adopted. By equating these mean values of the kinetic and potential energies there is obtained an expression for ν, the frequency of the constrained motion. According to Rayleigh's principle, this value of the frequency is always in excess of the natural frequency of the gravest mode of vibration. Moreover, it is usually found that almost any differentiable function satisfying the boundary conditions of the problem may be made the basis of a calculation yielding a close approximation to the fundamental frequency.

Although these results follow easily enough from the Lagrangian method of treating the oscillations of a system possessing only a finite number of degrees of freedom,[*] they do not appear to have been investigated in the far more important case of a continuous system. In this paper a closer examination of Rayleigh's principle is made by considering a series of successive approximations to the accurate solutions of the problems proposed, and a method is devised for obtaining an upper bound to the error involved in the approximate values of the frequency. Rayleigh's method is also extended to the calculation of the frequency of the first overtone, and analogous methods are given for more general problems of the computation of "characteristic numbers."

Before proceeding to develop the general theory of this paper, it is convenient to give a simple illustration of Rayleigh's method and to examine the nature of the difficulties which may be expected to follow the application of a rigorous mathematical treatment.

---

[*] Rayleigh, 'Scientific Papers,' vol. 1, p. 170.

Consider the transverse vibrations of a flexible string of variable density. Let the string have a length l and be stretched by a tension P. Let $\rho(x)$ be its mass per unit length at a distance x from one end. Then the differential equation for the transverse displacement $y(x, t)$ at a distance x from one end at time t is

$$\mathrm{P}\frac{\partial^2 y}{\partial x^2} = \rho(x)\frac{\partial^2 y}{\partial t^2},$$

and the boundary conditions are $y(x, t) = 0$ when $x = 0$ and when $x = l$ for all values of t. If the string is vibrating in a normal mode with frequency ν, we may write

$$y(x, t) = u(x)\sin 2\pi\nu t,$$

the function $u(x)$, which is the amplitude of the oscillations at any point, satisfying the equation

$$u''(x) + \lambda\,\rho(x)\,u(x) = 0, \qquad (1.1)$$

where $\lambda\mathrm{P} = 4\pi^2\nu^2$.

The values of ν for which the equation (1.1) possesses a continuous solution vanishing when $x = 0$ and when $x = l$ are the natural frequencies of the string. Let it be assumed that these values may be arranged in a discrete series of positive increasing terms

$$0 < \nu_0 < \nu_1 < \nu_2 < \ldots,$$

and that the functions, $\phi_0(x)$, $\phi_1(x)$, which represent the amplitude of the fundamental vibration and the first overtone, have, respectively, no zero and one zero in the interval, $0 < x < l$.

Let $f(x)$ be any function which possesses a continuous first derivative when $0 \leqslant x \leqslant l$, and which vanishes at the ends of this interval, then in virtue of the equations

$$\phi_0''(x) + \lambda_0\rho(x)\,\phi_0(x) = 0,$$

$$\lambda_0\mathrm{P} = 4\pi^2\nu_0^2,$$

we may form Picard's identity*

$$\frac{d}{dx}\left\{\frac{f^2(x)\,\phi_0'(x)}{\phi_0(x)}\right\} \equiv -\left\{\frac{f'(x)\,\phi_0(x) - f(x)\,\phi_0'(x)}{\phi_0(x)}\right\}^2 + f'^2(x) - \lambda_0\rho(x)f^2(x).$$

Integrating this identity from $x = 0$ to $x = l$, we find that

$$\int_0^l f'^2(x)\,dx - \lambda_0\int_0^l \rho(x)f^2(x)\,dx \geqslant \left[\frac{f^2(x)\,\phi_0'(x)}{\phi_0(x)}\right]_0^l.$$

---

* 'Traité d'Analyse,' vol. 3, p. 114 (1896).

The right-hand side of this inequality vanishes in virtue of the boundary conditions satisfied by the functions $\phi_0(x)$ and $f(x)$, and hence

$$\int_0^l f'^2(x)\,dx \geqslant \lambda_0 \int_0^l \rho(x) f^2(x)\,dx. \tag{1.2}$$

Now if the string were constrained to vibrate in such a manner that $f(x)$ was the amplitude at any point and ν the frequency, then the mean kinetic energy would be

$$T = \pi^2\nu^2 \int_0^l \rho(x) f^2(x)\,dx,$$

and the mean potential energy would be

$$V = \tfrac{1}{2}P \int_0^l \{(1+f'^2(x))^{\frac{1}{2}} - 1\}\,dx \doteqdot \tfrac{1}{4}P \int_0^l f'^2(x)\,dx.$$

It now follows from the preceding inequality (1.2) that the value of ν, obtained by equating T and V, cannot be less than the true value of the fundamental frequency ν_0.

§ 2. *Extension to the First Overtone.*

The preceding argument indicates that any function possessing a continuous derivative and satisfying the boundary conditions can be used to obtain an upper bound to the fundamental frequency. This method can be extended so as to obtain an upper bound to the frequency of the first overtone if a similar function $f(x)$ can be constructed which is orthogonal to $\rho(x)\,\phi_0(x)$, *i.e.*, if

$$\int_0^l \rho(x)\,\phi_0(x) f(x)\,dx = 0. \tag{2.1}$$

We have already assumed that in the mode of vibration which gives rise to the first overtone there is just one node. Let this node be at the point $x = s$, so that $\phi_1(s) = 0$, $(0 < s < l)$. Let

$$g(x) = \phi_0(x) f(s) - \phi_0(s) f(x), \tag{2.2}$$

then $g(s) = 0$, and the function $g(x)/\phi_1(x)$ is continuous throughout the interval $0 \leqslant x \leqslant l$. Hence, in virtue of the equations

$$\phi_1''(x) + \lambda_1\rho(x)\,\phi_1(x) = 0,$$
$$\lambda_1 P = 4\pi^2\nu_1^2,$$

we may form the identity, analogous to that used by Picard,

$$\frac{d}{dx}\left\{\frac{g^2(x)\,\phi_1'(x)}{\phi_1(x)}\right\} \equiv -\left\{\frac{g'(x)\,\phi_1(x) - g(x)\,\phi_1'(x)}{\phi_1(x)}\right\}^2 + g'^2(x) - \lambda_1\rho(x)\,g^2(x).$$

Since $g(x)$ vanishes when $x = 0$ or l, we may, as before, deduce the inequality

$$\int_0^l g'^2(x)\,dx \geqslant \lambda_1 \int_0^l \rho(x)\, g^2(x)\,dx.$$

In this inequality we substitute the value of $g(x)$ given by equation (2.2) and then in virtue of equation (2.1) we find that

$$\begin{aligned}\phi_0^2(s)\int_0^l f'^2(x)\,dx - \lambda_1\phi_0^2(s)\int_0^l \rho(x) f^2(x)\,dx \\ \geqslant \lambda_1 f^2(s)\int_0^l \rho(x)\,\phi_0^2(x)\,dx - f^2(s)\int_0^l \phi_0'^2(x)\,dx \\ = (\lambda_1 - \lambda_0) f^2(s)\int_0^l \rho(x)\,\phi_0^2(x)\,dx \geqslant 0,\end{aligned}$$

since the density $\rho(x)$ is never negative and $\lambda_1 > \lambda_0$.
Hence

$$\int_0^l f'^2(x)\,dx \geqslant \lambda_1 \int_0^l \rho(x) f^2(x)\,dx.$$

Thus any function, orthogonal to $\rho(x)\phi_0(x)$ and possessing a continuous derivative, may be used to obtain an upper bound to the frequency of the first overtone. Such functions may be constructed in various ways. If $h(x)$ is any function possessing a continuous derivative and vanishing when $x = 0$ or l, and the function $\phi_0(x)$ or the corresponding frequency ν_0 are known, we may write

$$f(x) = h(x)\int_0^l \rho(x)\,\phi_0^2(x)\,dx - \phi_0(x)\int_0 \rho(x)\,\phi_0(x)\,h(x)\,dx,$$

or

$$f(x) = h(x) - \lambda_0(x/l)\int_0^l (x-s)\,h(s)\,\rho(s)\,ds + \lambda_0\int_0^x (x-s)\,h(s)\,\rho(s)\,ds,$$

thus securing a suitable function for the calculation of an upper bound to λ_1.

If it were possible to construct a function orthogonal to the first n functions, $\phi_0(x)$, $\phi_1(x)$, ... $\phi_{n-1}(x)$, which give the amplitudes of the vibrations of the first n harmonics, then the application of a method similar to that given above would provide an upper limit to the $(n+1)$th frequency ν_n; but the labour of constructing such functions renders the investigation of little value.

We now proceed to consider critically the arguments outlined above.

§ 3. *Line and Band Spectra.*

It was explicitly assumed in § 1 that the values of ν for which the equation (1.1) possessed a continuous solution formed a discrete series of positive terms,

$$0 < \nu_0 < \nu_1 < \nu_2 < \dots.$$

But it is not difficult to imagine the mass of the string distributed in such a way that the frequency of the vibrations is susceptible of continuous variation within certain bounds. Consider, for example, the vibrations of a string in which

$$\rho(x) = x^{-2},$$

ignoring the facts that this assumption makes the total mass and kinetic energy infinite. For all positive values of λ the equation

$$u''(x) + \lambda x^{-2} u(x) = 0$$

has a solution continuous in the interval $0 \leqslant x \leqslant l$, possessing continuous first and second derivatives in the half-open interval $0 < x \leqslant l$, and vanishing at both ends of this interval.

Three cases arise. If $0 < \lambda < \frac{1}{4}$, we write $\kappa = \sqrt{(\frac{1}{4} - \lambda)}$, and the solution is $u(x) = \mathrm{A}x^{\frac{1}{2}}(x^{\kappa} - x^{-\kappa})$, A being any constant. If $\lambda = \frac{1}{4}$, the solution is $u(x) = \mathrm{A}x^{\frac{1}{2}} \log x$. If $\frac{1}{4} < \lambda$, we write $\kappa = \sqrt{(\lambda - \frac{1}{4})}$, and the solution is $u(x) = \mathrm{A}x^{\frac{1}{2}} \sin(\kappa \log x)$; thus a note of any frequency whatsoever could be elicited from such a string.

Now although the nature of this particular problem and of the solutions proposed do violence to our physical concepts, nevertheless the possibilities which have just been illustrated must be remembered in formulating a purely analytic theory.

It is convenient to refer to the set of numbers, $\nu_0, \nu_1, \dots$, *i.e.*, the frequencies of the normal modes of vibration, as the " spectrum " of the string. If the set of numbers form a discrete series we have a " line spectrum " : if the frequency is susceptible of continuous variation we have a " band spectrum." Of course, both varieties of spectrum may occur in the same problem, as in the theory of the hydrogen atom according to Schroedinger's wave-mechanics.

Again, it follows at once from the equations

$$\int_0^l \phi_n'^2(s)\,dx = -\int_0^l \phi_n(x)\,\phi_n''(x)\,dx = \lambda_n \int_0^l \rho(x)\,\phi_n^2(x)\,dx$$

that all the real values of λ for which equation (1.1) possesses a continuous solution, vanishing when x is 0 or l, are positive numbers. But when the problem is extended by the introduction of more general boundary conditions this result is no longer necessarily true. For example, consider the equation

$$u''(x) + \lambda u(x) = 0$$

and the boundary conditions $u'(x) + u(x) = 0$, when $x = 0$ or l. Corresponding to the value $\lambda = -1$, there is a solution

$$u(x) = Ae^{-x}.$$

It will be shown in § 5 that the types of differential equations considered in this paper are such that the spectrum is always a line spectrum, and that the special (or characteristic) values of the parameter λ are always real positive numbers.

§ 4. *Multiple Solutions.*

It was implied, although not actually stated, in § 1 that all the solutions of the equation (1.1),

$$u''(x) + \lambda\rho(x)u(x) = 0,$$

which corresponded to a characteristic value of λ, differed among themselves only by numerical multipliers. That this attractive hypothesis is not necessarily true when we study the more general problems outlined in § 1 may be seen from a consideration of the simplest equation,

$$u''(x) + \lambda u(x) = 0, \tag{4.1}$$

associated with the periodic boundary conditions,

$$u(0) = u(l), \quad u'(0) = u'(l).$$

The smallest characteristic number is $\lambda_0 = 0$ and all the corresponding solutions of equation (4.1) are numerical multiples of the solution $\phi_0(x) = 1$; but corresponding to any other characteristic number $\lambda_n = 4\pi^2n^2l^{-2}$ (where n is a positive integer) there are two linearly independent solutions,

$${}_1\phi_n(x) = A \sin 2\pi nx/l, \qquad {}_2\phi_n(x) = B \cos 2\pi nx/l. \tag{4.2}$$

It might be thought from this example that although there may be multiple solutions associated with the higher characteristic numbers, nevertheless the first, or smallest, characteristic number cannot have two linearly independent associated solutions. Unfortunately this is not the case, for an example can be constructed in which the first characteristic number yields multiple solutions while every other characteristic number yields a simple solution.

Consider the same equation as before,

$$u''(x) + \lambda u(x) = 0,$$

associated with the boundary conditions,

$$u(l) - u(0) = lu'(l) = lu'(0),$$

which make the mean gradient of the function $u(x)$ equal to its gradient at either end of the interval. There are two linearly independent solutions associated with the first characteristic number $\lambda_0 = 0$, viz.,

$$ {}_1\phi_0(x) = \mathrm{A}, \qquad {}_2\phi_0(x) = \mathrm{B}x, $$

but none of the higher characteristic numbers can possess multiple solutions.

The methods to be given later will take explicit cognisance of the existence of multiple solutions and will furnish the appropriate modifications of Rayleigh's method.

§ 5. *The Specification of the Vibrating System.*

Although the subsequent argument does not depend upon any dynamical principles but is based on mathematical analysis alone, nevertheless it is convenient to make use of the terminology of dynamics in naming the various functions which rise in the course of the work.

Consider a vibrating bar of length l. At a distance x from one end let $p(x)$ be the flexural rigidity, $q(x)$ the longitudinal tension, and $\sigma(x)$ the mass per unit length. If $r(x)$ is the restoring force per unit length per unit displacement, then the equation for the displacement $y(x, t)$ is

$$ -\{p(x)\,y''(x,t)\}'' + \{q(x)\,y'(x,t)\}' - r(x)\,y(x,t) = \mathrm{L}\{y(x,t)\} $$
$$ = \sigma(x)\,\ddot{y}(x,t). \quad (5.1) $$

When the bar vibrates in a normal mode with frequency ν, we may write

$$ y(x,t) = u(x)\sin 2\pi\nu t, $$

and the function $u(x)$, which measures the amplitude of the vibrations, will satisfy the equation

$$ \mathrm{L}(u) + \lambda\sigma(x)\,u(x) = 0, \quad (5.2) $$

where

$$ \lambda = 4\pi^2\nu^2. $$

The mean potential energy of the bar is

$$ \mathrm{V}(u) = \tfrac{1}{4}\int_0^l \{p(x)\,u''^2(x) + q(x)\,u'^2(x) + r(x)\,u^2(x)\}\,dx, \quad (5.31) $$

and the mean kinetic energy is

$$ \mathrm{T}(u) = \pi^2\nu^2\int_0^l \sigma(x)\,u^2(x)\,dx. \quad (5.32) $$

If the bar is subjected to a steady transverse force distributed with intensity $\mathrm{F}(x)$ per unit length, then the equation for the deflection $y(x)$ is

$$ \mathrm{L}(y) + \mathrm{F}(x) = 0, \quad (5.4) $$

and the work done by this force is $\tfrac{1}{2}\int_0^l y(x)\,\mathrm{F}(x)\,dx = -\tfrac{1}{2}\int_0^l y(x)\,\mathrm{L}(y)\,dx$. Let

$y_1(x)$, $y_2(x)$ be the deflections produced by forces distributed with intensities $F_1(x)$ and $F_2(x)$, then $y_1(x) + y_2(x)$ is the deflection produced by a force distributed with intensity $F_1(x) + F_2(x)$. This deflection may be produced by applying the two systems of forces successively. If the force of intensity $F_1(x)$ is first applied, the work done is

$$-\int_0^l \{\tfrac{1}{2}y_1(x)\,L(y_1) + y_2(x)\,L(y_1) + \tfrac{1}{2}y_2(x)\,L(y_2)\}\,dx:$$

if the order of operations is reversed the work done is

$$-\int_0 \{\tfrac{1}{2}y_2(x)\,L(y_2) + y_1(x)\,L(y_2) + \tfrac{1}{2}y_1(x)\,L(y_1)\}\,dx.$$

If the same amount of work is done in both cases, we must have the relation

$$\int_0^l y_1(x)\,L(y_2)\,dx = \int_0^l y_2(x)\,L(y_1)\,dx. \tag{5.5}$$

The differential expression $L(u)$ is easily seen to be self-adjoint from the form of equation (5.1). Hence the expression $y_1(x)\,L(y_2) - y_2(x)\,L(y_1)$ is a perfect differential and the integral of this expression will depend only upon the values of $y_1(x)$, $y_2(x)$ and their first three derivatives at the points $x = 0$, $x = l$. Hence the equation (5.5) can be satisfied only in virtue of the boundary conditions which are fulfilled by any function representing a possible deflection.

Moreover, if $2V(y)$, the potential energy stored in the bar when deflected by a force distributed with intensity $F(x)$, is equal to the work done by this force, then we must also have the relation

$$\int_0^l y(x)\,F(x)\,dx = -\int_0^l y(x)\,L(y)\,dx = 4V(y), \tag{5.6}$$

and again it may be shown that this equation can be satisfied only in virtue of the boundary conditions.

These physical considerations prepare the way for the following formal specification of the mathematical problem. The differential system which we are studying consists of a fourth order linear homogeneous self-adjoint equation (5.2), together with a set of four linear homogeneous boundary conditions "M," of the form

$$\sum_{n=0}^{3} \{a_{mn}y^{(n)}(0) + b_{mn}y^{(n)}(l)\} = 0, \qquad (m = 1, 2, 3, 4),$$

a_{mn}, b_{mn} being constants. The coefficients which occur in equation (5.2) are subject to these further restrictions: In the interval, $0 \leqslant x \leqslant l$, $\sigma(x)$ and $r(x)$ are never negative and are everywhere finite and continuous, $q(x)$ is never negative and is finite and continuous together with its first derivative, $p(x)$ is always positive and is finite and continuous together with its first and second derivatives. The boundary conditions are also restricted to be sufficient to ensure that any pair of functions, $y_1(x)$ and $y_2(x)$, which satisfy them, shall also satisfy Green's formula (5.5) and the energy relation (5.6). Finally, we suppose that the equation $\mathrm{L}(y) = 0$ does not possess any continuous solution satisfying the boundary conditions "M" other than $y = 0$. (This implies that in the absence of any external force the bar remains undeflected everywhere.) It is required to determine the values of λ for which this differential system possesses continuous solutions.

This problem may also be treated as an integral equation by the introduction of the Green's function, $\mathrm{G}(x, s)$. This function determines the steady deflection produced at the point x by a unit steady concentrated force at s. It is finite and continuous together with its first and second derivatives in the interval, $0 \leqslant x \leqslant l$. In the half-open intervals $0 \leqslant x < s$, $s < x \leqslant l$ it possesses continuous third and fourth derivatives, and satisfies the equation $\mathrm{L}(y) = 0$. The discontinuity in the third derivative at $x = s$ is given by the equation

$$p(s)\,\mathrm{G}'''(s + 0, s) - p(s)\,\mathrm{G}'''(s - 0, s) = 1.$$

Finally, $\mathrm{G}(x, s)$ also satisfies the boundary conditions "M."

The existence of a unique function satisfying these conditions is ensured by the restrictions which have been placed upon the differential system we are studying. The principal property of this function is that it allows the continuous solution of equation (5.4) which fulfils the boundary conditions "M" to be written in the form

$$h(x) = \int_0^l \mathrm{G}(x, s)\,\mathrm{F}(s)\,ds. \tag{5.7}$$

Now if we consider the bar subjected to a periodic force distributed with intensity $\mathrm{F}(x) \sin 2\pi\nu t$ per unit length, the vibrations of the same period elicited by this force will have an amplitude $u(x, \lambda)$ which satisfies the equation

$$\mathrm{L}(u) + \lambda\sigma(x)\,u(x, \lambda) + \mathrm{F}(x) = 0. \tag{5.8}$$

It now follows from equation (5.7) that $u(x, \lambda)$ satisfies the integral equation

$$u(x, \lambda) = h(x) + \lambda \int_0^l \sigma(s)\,\mathrm{G}(x, s)\,u(s, \lambda)\,ds, \tag{5.9}$$

and that the values of λ required are the characteristic numbers of this equation.

It may now be shown, by means of classical theorems, that since $\sigma(s)\,\mathrm{G}(x, s)$ is bounded, the characteristic numbers are isolated, so that there is no "band spectrum." Also the Green's function $G(x, s)$ is symmetric and hence the characteristic numbers are all real. Moreover, the Green's function is "closed," *i.e.*, the integral $\int_0^l G(x, s) F(s)\, ds$ cannot vanish for all values of x in the interval, $0 \leqslant x \leqslant l$, unless $F(x)$ is also zero in the same interval. Hence the number of characteristic numbers is infinite. Finally, we note that the integrals

$$2V(u) = \tfrac{1}{2}\int_0^l F(x)\, u(x)\, dx = \tfrac{1}{2}\int_0^l\int_0^l F(x)\, F(s)\, G(x, s)\, dx\, ds$$

(which give the potential energy stored in the bar by the action of a steady force distributed with intensity $F(x)$) are all positive unless $F(x) \equiv 0$. Hence $G(x, s)$ is a function of positive definite type, and, therefore, all the characteristic numbers are positive.

§ 6. *The Characteristic Numbers.*

The solution of the differential equation (5.8) for the amplitude $u(x, \lambda)$ of the forced vibrations may be expressed in terms of $G(x, s;\ \lambda)$, the Green's function of the equation (5.2) for the free oscillations of the bar. This function possesses analytical properties similar to those of $G(x, s)$, except that in the intervals $0 \leqslant x < s$, $s < x \leqslant l$ it satisfies equation (5.2); and it may be given an analogous physical interpretation as the amplitude of the oscillation produced at the point x by a concentrated periodic force $\sin 2\pi\nu t$ at s. The nature of the mathmetical problem specified in the last section shows that $G(x, s;\ \lambda)$ is an analytic function of λ except when λ is equal to a characteristic number. Hence, the function

$$u(x, \lambda) = \int_0^l G(x, s;\ \lambda)\, F(s)\, ds \quad . \qquad (6.1)$$

is also analytic, except perhaps when λ is equal to a characteristic number; it satisfies equation (5.8) and fulfils the boundary conditions "M."

The First Characteristic Number.—The mean value of the work done upon the bar by the action of the external forces is $\tfrac{1}{4}W(\lambda)$, where

$$W(\lambda) = \int_0^l u(x, \lambda)\, F(x)\, dx. \qquad (6.2)$$

This function can only fail to be analytic when λ is equal to a characteristic number. When $\lambda = 0$, *i.e.*, when a steady force is applied to the bar, $W(0) = V(u)$ and is positive unless $F(x) \equiv 0$. But for other values of λ

$$\begin{aligned} W(\lambda) &= -\int_0^l u(x, \lambda)\, L\{u(x, \lambda)\}\, dx - \lambda\int_0^l \sigma(x)\, u^2(x, \lambda)\, dx \\ &= 4V\{u(x, \lambda)\} - 4T\{u(x, \lambda)\}, \end{aligned}$$

and this expression is not necessarily positive. We shall show, however, that at any point at which $dW(\lambda)/d\lambda$ exists it is essentially positive.

Let $u_1(x, \lambda)$ and $u_2(x, \lambda)$ be the solutions of equation (5.8) fulfilling the boundary conditions "M" when λ has the real values λ_1 and λ_2, which are not characteristic numbers themselves and are not separated by any characteristic number. Then it follows from an application of Green's formula (5.5) that

$$\int_0^l u(x, \lambda_1) \mathrm{L}\{u(x, \lambda_2)\}\, dx = \int_0^l u(x, \lambda_2) \mathrm{L}\{u(x, \lambda_1)\}\, dx,$$

whence

$$\mathrm{W}(\lambda_1) - \mathrm{W}(\lambda_2) = (\lambda_1 - \lambda_2) \int_0^l \sigma(x)\, u(x, \lambda_1)\, u(x, \lambda_2)\, dx.$$

As $\lambda_2 \to \lambda_1$ the integral on the right-hand side of this equation tends towards the positive quantity $\int_0^l \sigma(x)\, u^2(x, \lambda_1)\, dx$, and therefore

$$\frac{d\mathrm{W}(\lambda)}{d\lambda} \doteq \lim_{\lambda_2 \to \lambda_1} \frac{\mathrm{W}(\lambda_2) - \mathrm{W}(\lambda_1)}{\lambda_2 - \lambda_1} = \int_0^l \sigma(x)\, u^2(x, \lambda_1)\, dx > 0.$$

This property of the energy function $\mathrm{W}(\lambda)$ furnishes an immediate proof of Rayleigh's principle. Let $h(x)$ be any function which is continuous together with its first four derivatives in the interval $0 \leqslant x \leqslant l$, and which satisfies the boundary conditions "M." Then this function determines a possible distribution of the amplitudes of the vibrations of the bar when excited by some external force. If $\sqrt{\mu}/2\pi$ is the frequency of these vibrations, their mean kinetic energy is

$$\mathrm{T}(h) = \tfrac{1}{4}\mu \int_0^l \sigma(x)\, h^2(x)\, dx, \tag{6.3}$$

while their mean potential energy is

$$\mathrm{V}(h) = -\tfrac{1}{4} \int_0^l h(x)\, \mathrm{L}\{h(x)\}\, dx. \tag{6.4}$$

If $\mathrm{F}(x) \sin \sqrt{\mu} t$ is the intensity per unit length of the external force exciting these oscillations,

$$\mathrm{F}(x) = -\mathrm{L}\{h(x)\} - \mu\sigma(x)\, h(x). \tag{6.5}$$

Now let $u(x, \lambda)$ denote the amplitude of the vibrations produced when the frequency is changed to $\sqrt{\lambda}/2\pi$. Then $u(x, \lambda)$ will satisfy equation (5.8), and the energy function associated with this system will be

$$\mathrm{W}(\lambda) = \int_0^l u(x, \lambda)\, \mathrm{F}(x)\, dx = -\int_0^l u(x, \lambda)\, \mathrm{L}\{h(x)\} - \mu \int_0^l \sigma(x)\, h(x)\, u(x, \lambda)\, dx.$$

In particular, the energy function for the original frequency $\sqrt{\mu}/2\pi$ is

$$W(\mu) = 4V(h) - 4T(h).$$

Applying Rayleigh's principle we choose μ so that $V(h) = T(h)$ and then we find that $W(\mu) = 0$. But the energy function is positive when $\lambda = 0$, and therefore cannot possess a positive derivative at all points in the real interval $0 \leqslant \lambda \leqslant \mu$. Hence there must be at least one characteristic number λ_0 less than, or equal to, μ.

The method by which we have obtained an upper bound to the fundamental frequency of the bar may be extended to the determination of an upper bound to the frequency of the first overtone. It has been indicated above that if λ is not equal to a characteristic number, then $u(x, \lambda)$ the amplitude of the vibrations elicited by an applied force of intensity $F(x) \sin \sqrt{\lambda}t$ is uniquely determined by equation (6.1). But when λ equals a characteristic number λ_0 the period of the impressed force is equal to one of the natural periods of the bar, and the amplitude of the oscillations will in general increase until the bar is ruptured, since no provision has been made for damping the vibrations. We shall now show that this resonance may be suppressed by distributing the external force in such a way that it does no work in eliciting from the bar any of the natural vibrations of period equal to that of the applied force.

The Second Characteristic Number.—The mode in which the Green's function $G(x, s; \lambda)$ originates shows that it is the ratio of two integral functions of λ. Accordingly it can possess only accidental singularities, and since the original Green's function $G(x, s)$ was symmetric in x and s, it follows that the singularities of $G(x, s; \lambda)$ are all simple poles. If $E(x, s; \lambda_0)$ is the residue of $G(x, s; \lambda)$ at the pole $\lambda = \lambda_0$, then Hilbert's function

$$H(x, s; \lambda) = G(x, s; \lambda) - \frac{E(x, s; \lambda_0)}{\lambda - \lambda_0}$$

will be analytic throughout the region $|\lambda| < |\lambda_1|$, where λ_1 is the characteristic number next in order of magnitude to λ_0. $E(x, s; \lambda_0)$ will be a linear combination of the characteristic functions associated with λ_0, and will be symmetrical in x and s.

The number, κ, of linearly independent characteristic functions associated with λ_0 cannot exceed four; and if $F(x)$ is orthogonal to each member of a set of κ such functions, it will also be orthogonal to $E(x, s; \lambda_0)$, *i.e.*,

$$\int_0^l F(x)\, E(x, s; \lambda_0)\, dx = 0. \tag{6.6}$$

It follows that when these conditions are fulfilled an applied force of intensity $F(x) \sin \sqrt{\lambda_0} \,.\, t$ would transfer no energy to the bar if it were vibrating in a fundamental mode. The amplitude of the vibrations elicited by such a force when the frequency is changed to $\sqrt{\lambda}/2\pi$ is given by the integral

$$u(x, \lambda) = \int_0^l G(x, s;\ \lambda_0) F(s)\, ds = \int_0^l H(x, s;\ \lambda_0) F(s)\, ds,$$

in virtue of equation (6.6), if $\lambda \neq \lambda_0$. The latter integral is analytic in the region $|\lambda| < |\lambda_1|$, and at every point in this region, with the possible exception of $\lambda = \lambda_0$, it satisfies equation (5.2), fulfils the boundary conditions "M," and is orthogonal to $\sigma(x) E(x, s;\ \lambda_0)$, for

$$\int_0^l \sigma(x) u(x, \lambda) E(x, s; \lambda_0)\, dx = \int_0^l \sigma(x) H(x, s; \lambda)\, dx \int_0^l E(x, s; \lambda_0) F(s)\, ds = 0.$$

It follows that the function $u(x, \lambda)$ defined by this integral will continue to possess the same properties even when $\lambda = \lambda_0$.

The energy function associated with these forced vibrations will be analytic when $|\lambda| < |\lambda_1|$, and will possess a positive derivative when λ is real and less than λ_1. In order to extend Rayleigh's principle to the determination of an upper bound to λ_1, we repeat the argument on p. 286, now choosing $h(x)$, the amplitude of the oscillations, to be orthogonal to the product of $\sigma(x)$ by any characteristic function, $\phi_0(x)$, associated with λ_0. It then follows from Green's formula (5.5) that

$$\int_0^l F(x)\, \phi_0(x)\, dx = -\mu \int_0^l \sigma(x)\, h(x)\, \phi_0(x)\, dx - \int_0^l \phi_0(x)\, L(h)\, dx$$

$$= -\int_0^l h(x)\, L(\phi_0)\, dx = \lambda_0 \int_0^l \sigma(x)\, h(x)\, \phi_0(x)\, dx = 0.$$

Hence $F(x)$ will be orthogonal to $\phi_0(x)$, and the equation

$$L(u) + \lambda\sigma(x) u(x, \lambda) = -F(x) = L(h) + \mu\sigma(x)\, h(x)$$

will possess a solution $u(x, \lambda)$ analytic in the interval $0 \leqslant \lambda < \lambda_1$. If μ is chosen as before so that $V(h) = T(h)$, it follows that the energy function associated with $u(x, \lambda)$ will vanish when $\lambda = \mu$, although it is analytic and steadily increasing when $0 \leqslant \lambda < \lambda_1$. Hence it ceases to be analytic at some point in the interval $\lambda_0 < \lambda \leqslant \mu$. Consequently the second characteristic number λ_1 cannot exceed the value of μ determined by Rayleigh's method.

§ 7. *Successive Approximations.*

The Generating Function.—It will be seen from the preceding section that in order to use Rayleigh's method of obtaining an upper bound to λ_0 we require function $h(x)$ which is continuous, together with its first four derivatives when $0 < x < l$, and such that $\sigma(x) h(x)$ is *not* orthogonal to all the characteristic functions associated with λ_0. If $G(x, 0)$ and $G(x, l)$ are not zero, they satisfy these conditions. If $G(x, s)$ vanishes when $s = 0$ and has a zero of multiplicity m, we write

$$f_0(x) = \lim_{s \to 0} s^{-m} G(x, s), \qquad f_0(s) = \lim_{s \to 0} x^{-m} G(x, s),$$

and similarly if $G(x, l) = 0$. If $\phi_0(x)$ is any characteristic function associated with λ_0,

$$\phi_0(x) = \lambda_0 \int_0^l \sigma(s) G(x, s) \phi_0(s) \, ds.$$

Since $\phi_0(x)$ fulfils the same boundary conditions as $G(x, s)$, it follows that $\phi_0(x)$ also has a zero of multiplicity m when $x = 0$, and hence

$$\int_0^l \sigma(x) f_0(x) \phi_0(x) \, dx = \lim_{s \to 0} s^{-m} \int_0^l \sigma(x) G(x, s) \phi_0(x) \, dx$$

$$= \frac{1}{\lambda_0} \lim_{s \to 0} s^{-m} \phi_0(s), \tag{7.1}$$

which does not vanish.

Since $G(x, 0)$ is continuous and possesses continuous derivatives when $0 < x \leqslant l$, it follows that $f_0(x)$ fulfils all the conditions stipulated above.

The Sequence of Approximations.—We shall now show that the approximation to λ_0 obtained by applying Rayleigh's method to $f_0(x)$ may be improved by applying the same method to any function of the sequence derived from $f_0(x)$ by Picard's method.[*] This sequence is defined by the recurrence relation

$$f_{n+1}(x) = \int_0^l \sigma(s) G(x, s) f_n(s) \, ds, \qquad (n = 0, 1, 2, \ldots). \tag{7.2}$$

It follows by successive applications of Green's formula (5.5) that all the integrals

$$\int_0^l \sigma(x) f_m(x) f_{n-m}(x) \, dx, \qquad (m = 0, 1, 2, \ldots n)$$

have the same value, which we shall denote by a_n. By integrating the function

$$\sigma(x) \sigma(s) \{f_{n-1}(x) G(y, s) - f_{n-1}(y) G(x, s)\}^2$$

---

[*] *Loc. cit.*, p. 105.

over the square $0 \leqslant x, y \leqslant l$, we may now deduce that

$$f_n^2(s) \leqslant a_{2n-2} \int_0^l \sigma(x)\, G^2(x, s)\, dx, \tag{7.3}$$

—an inequality which is required later.

If the bar is constrained to vibrate with frequency ν and amplitude

$$u(x) = \alpha f_n(x) + \beta f_{n+1}(x),$$

(α and β constants), its mean potential energy will be

$$\begin{aligned} V(u) &= -\tfrac{1}{4}\int_0^l u(x)\, L(u)\, dx \\ &= \tfrac{1}{4}\int_0^l \sigma(x)\, \{\alpha f_{n-1}(x) + \beta f_n(x)\}\{\alpha f_n(x) + \beta f_{n+1}(x)\}\, dx, \end{aligned}$$

and its mean kinetic energy will be

$$T(u) = \pi^2\nu^2 \int_0^l \sigma(x)\, u^2(x)\, dx = \pi^2\nu^2 \int_0^l \sigma(x)\, \{\alpha f_n(x) + \beta f_{n+1}(x)\}^2\, dx.$$

Since both of these quantities are essentially positive, we obtain the inequalities

$$a_{2n-1}\alpha^2 + 2a_{2n}\alpha\beta + a_{2n+1}\beta^2 > 0,$$

and

$$a_{2n}\alpha^2 + 2a_{2n+1}\alpha\beta + a_{2n+2}\beta^2 > 0;$$

which imply that

$$a_{2n-1}a_{2n+2} > a_{2n}^2, \qquad a_{2n}a_{2n+2} > a_{2n+1}^2.$$

Rayleigh's method applied to $f_n(x)$, yields

$$\mu_{2n-1} = a_{2n-1}/a_{2n}$$

as an upper bound to λ_0. If we define μ_{2n} as the ratio a_{2n}/a_{2n+1}, we find from the preceding inequalities that

$$\mu_{2n-1} > \mu_{2n} > \mu_{2n+1}.$$

Hence the sequence of approximations, $\mu_0, \mu_1, \mu_2, \ldots, \mu_n, \mu_{n+1}, \ldots$, steadily decreases, each member, of course, remaining superior to λ_0. Therefore this sequence converges to a unique limit $\mu \geqslant \lambda_0$.

Series Solutions of the Equation for Forced Vibrations (5.8).—It follows from the inequality (7.3) that the series

$$u(x, \lambda) = f_1(x) + \lambda f_2(x) + \ldots + \lambda^{n-1} f_n(x) + \ldots \tag{7.4}$$

is absolutely and uniformly convergent if the series

$$\sqrt{a_0} + \lambda\sqrt{a_1} + \ldots + \lambda^{n-1}\sqrt{a_{2n-2}} + \ldots \tag{7.5}$$

converges. Now as $n \rightarrow \infty$, $\sqrt{(a_{2n-2}/a_{2n})} = \sqrt{(\mu_{2n-2}\mu_{2n-1})} \rightarrow \mu$, whence we deduce that the radius of convergence of the series (7.5) is $\lambda = \mu$. Hence, when $\lambda < \mu$, the series

$$-\sigma(x)\{f_0(x) + \lambda f_1(x) + \dots + \lambda^n f_n(x) + \dots\}$$

derived from (7.4) also converges absolutely and uniformly to $L\{u(x, \lambda)\}$. Therefore the series (7.4) provides a solution of the equation for forced vibrations (5.8). Each term in the series satisfies the boundary conditions " M " and hence $u(x, \lambda)$ represents the amplitude of the oscillations produced by an external force of intensity $f_0(x) \sin\sqrt{\lambda} \,.\, t$ per unit mass.

If $\mu > \lambda_0$, the series (7.4) will provide a solution even when $\lambda = \lambda_0$. But it follows from Green's formula (5.5) that, if $\phi_0(x)$ is any characteristic function associated with λ_0, then

$$\int_0^l \sigma(x) f_0(x) \phi_0(x) dx = -\int_0^l \phi_0(x) L\{u(x, \lambda_0)\} dx - \lambda_0 \int_0^l \sigma(x) \phi_0(x) u(x, \lambda_0) dx = 0.$$

Since this result is in contradiction with equation (7.1) we conclude that $\mu = \lambda_0$.

Solution in the Case of Suppressed Resonance.

Let

$$\Psi(s) = \int_0^l \sigma(x) f_0(x) E(x, s;\ \lambda_0) dx$$

and

$$g_0(x) = f_0(x) - \Psi(x).$$

Then, since $E(x, s;\ \lambda_0)$—the residue of $G(x, s;\ \lambda_0)$ at $\lambda = \lambda_0$—may be written in the form

$$E(x, s;\ \lambda_0) = {}_1\phi_0(x) \cdot {}_1\phi_0(s) + \dots + {}_\kappa\phi_0(x) \cdot {}_\kappa\phi_0(s),$$

where ${}_1\phi_0(x), \dots {}_\kappa\phi_0(x)$ form any set of κ normal and orthogonal characteristic functions for $\lambda = \lambda_0$: therefore $g_0(x)$ is orthogonal to all the characteristic functions associated with λ_0.

With $g_0(x)$ as a new generating function we form the sequence of functions $g_1(x), g_2(x), \dots g_n(x), \dots$ defined by the recurrence relation

$$g_{n+1}(x) = \int_0^l \sigma(s) G(x, s) g_n(s) ds = f_{n+1}(x) - \lambda_0^{-n-1} \Psi(x), \quad (n = 0, 1, 2, \dots).$$

We also introduce a sequence of constants defined by the equation

$$b_n = \int_0^l \sigma(s) g_m(x) g_{n-m}(x) dx = a_n - c\lambda_0^{-n}, \quad \begin{matrix}(m = 0, 1, 2, \dots n) \\ (n = 0, 1, 2, \dots)\end{matrix} \tag{7.6}$$

where

$$c = \int_0^l \sigma(x) f_0(x) \Psi(x) dx.$$

We may now show as before that the sequence $\nu_0, \nu_1, \nu_2, \ldots, \nu_n, \nu_{n+1} \ldots,$ where

$$\nu_n = b_n/b_{n+1},$$

decreases steadily to the unique limit λ_1, while the series

$$u(x, \lambda) = g_1(x) + \lambda g_2(x) + \ldots + \lambda^{n-1} g_n(x) + \ldots$$

is absolutely and uniformly convergent when $|\lambda| < |\lambda_1|$, and determines the amplitude of the oscillations produced by an external force of intensity $g_0(x) \sin \sqrt{\lambda} . t$ per unit mass.

§ 8. *Upper and Lower Bounds to the First Characteristic Number.*

It follows from equation (7.6) that

$$a_n - \lambda_0 a_{n+1} = b_n - \lambda_0 b_{n+1},$$

whence we deduce that

$$\frac{\mu_n - \lambda_0}{\mu_{n+1} - \lambda_0} . \mu_{n+1} = \frac{a_n - \lambda_0 a_{n+1}}{a_{n+1} - \lambda_0 a_{n+2}} = \frac{b_n - \lambda_0 b_{n+1}}{b_{n+1} - \lambda_0 b_{n+2}} = \frac{\nu_n - \lambda_0}{\nu_{n+1} - \lambda_0} \nu_{n+1}.$$

Now the last expression exceeds ν_{n+1}, since $\nu_n > \nu_{n+1}$. And since $\nu_{n+1} > \lambda_1$, therefore $(\mu_n - \lambda_0) \mu_{n+1} > (\mu_{n+1} - \lambda_0) \lambda_1$, and

$$\mu_{n+1} - \lambda_0 < (\mu_n - \mu_{n+1})/(\lambda_1/\mu_{n+1} - 1). \tag{8.1}$$

To give this inequality a more useful form we note that, by a well-known theorem, the integral

$$\Gamma = \int_0^l \int_0^l \sigma(x) \sigma(s) G^2(x, s) \, dx \, ds \tag{8.2}$$

has a value exceeding $\kappa_0 \lambda_0^{-2} + \kappa_1 \lambda_1^{-2}$, where κ_0 and κ_1 are the numbers of linearly independent characteristic functions associated with λ_0 and λ_1. Hence, since $\mu_n \rightarrow \lambda_0$ as $n \rightarrow \infty$, the inequality

$$\mu_n^2 \Gamma < \kappa_0 + \kappa_1$$

will be satisfied for all terms of the sequence $\{\mu_n\}$, except, perhaps, a finite number at the beginning. We suppose that by actual calculation of Γ and μ_n, we have found a number J_n, such that

$$\mu_n^2 \Gamma = \kappa_0 + \frac{\kappa_1}{J_n^2}, \quad \text{where} \quad J_n > 1. \tag{8.3}$$

Then we find that

$$\frac{\kappa_1 \mu_n^2}{\lambda_1^2} < \mu_n^2 \Gamma - \frac{\kappa_0 \mu_n^2}{\lambda_0^2} = \frac{\kappa_1}{J_n^2} - \kappa_0 \left\{ \frac{\mu_n^2}{\lambda_0^2} - 1 \right\} < \frac{\kappa_1}{J_n^2}.$$

Hence $\lambda_1/\mu_n > J_n$, and

$$\mu_n - \lambda_0 < (\mu_{n-1} - \mu_n)/(J_n - 1). \tag{8.4}$$

This inequality provides an upper bound to the error involved in taking μ_n to be an approximation to λ_0.

We may deduce from equation (8.1) that the difference $(\mu_n - \lambda_0)$ can never be as small as $(\mu_{n-1} - \mu_n)/(\lambda_1/\lambda_0 - 1)$, and hence the success of Rayleigh's method in providing a good approximation to λ_0 depends upon the magnitude of the ratio λ_1/λ_0. In the majority of practical applications which have so far been examined, this ratio has been always greater than two, so that the sequence $\{\mu_n\}$ has converged with fair rapidity. In a subsequent communication I hope to give some illustrations of the facility with which Rayleigh's method can be applied to the determination of the smallest characteristic number of a differential system with close and *known* limits of approximation. It will be recognised that although this paper has dealt with equations of the fourth order only, the methods employed may be extended at once to equations of higher order.

On Phenomena occurring in the Condensation of Molecular Streams on Surfaces.

By J. D. Cockcroft, B.A.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received March 29, 1928.)

[Plates 2 and 3.]

(1) *Introduction.*

R. W. Wood* first investigated the phenomena occurring when a stream of metallic vapour is allowed to impinge on a surface of glass or other material, the vessel in which the experiment takes place being evacuated. Under these conditions the rate of condensation depends very largely on the temperature of the surface.

Wood showed that for temperatures below −140° C. the probability of a mercury atom adhering to the surface, after one impact, was unity; no atoms were reflected; whilst for higher temperatures only a proportion of the atoms could be considered to adhere after the first impact. Similar phenomena were observed for other metals.

Knudsen† showed that the probability ε of an atom being adsorbed after one collision is a decreasing function of temperature and that there exists a "critical temperature" where ε changes suddenly, separating the regions of very rapid condensation from those of slow condensation.

---

* 'Phil. Mag.,' vol. 30, p. 300 (1915).

† 'Ann. de Phys.,' vol. 50, p. 472 (1916).

Knudsen, however, only determined this critical temperature very approximately, showing, for example, that for zinc, cadmium, and mercury it lay between —183° and —78°.

Further experiments have been made by Chariton and Semenoff* and Estermann† over narrow ranges of temperatures, showing that this critical temperature decreases with decreasing stream intensity. It is of interest, however, to investigate the phenomenon over a much wider range of temperatures and by more accurate methods, for two reasons. First, in experiments with molecular streams where it is desired to derive information as to the intensity of the incident stream from the thickness of the deposit found, it is of importance to know what proportion of the molecules condense under given conditions. Secondly, the experiments might be expected to throw a great deal of light on the nature and magnitude of the forces between the surface and the incident molecules.

The phenomena were studied in three separate series of experiments. In the first, an atomic stream of approximately constant density was allowed to fall on a condensing surface in which a temperature gradient exists and the region of deposition observed. In the second series, an atomic stream was used whose density varied from point to point of a condensing surface of constant temperature. In the third series, the effects of different surfaces and different surface conditions were investigated.

(2) *Experiments with Constant Stream Densities.*

The first apparatus used was practically identical with that of Chariton and Semenoff (*loc. cit.*) and is shown diagrammatically in fig. 1. A platinum filament on which cadmium had been deposited electrolytically was stretched parallel to a copper plate on which was cemented a sheet of mica. The apparatus was placed in a glass tube and exhausted to a pressure of below 10^{-5} mm. The tube was immersed in liquid air and a uniform temperature gradient of about two degrees per centimetre maintained in the copper plate by making suitable thermal contact between the bottom of the plate and the liquid air outside, the mean temperature of the plate being between —60° C. and —120° C. in different experiments. The filament was heated electrically and the cadmium evaporated in a stream whose density was approximately constant along the length of the plate but decreased across its breadth from the centre to the sides. As in the experiments of Chariton and Semenoff it was found that a deposit occurred below a certain point on the plate, the boundary of the deposit being sharply defined to 1/10th of a degree and having a slight curvature as shown on fig. 1. The temperature of the plate at the boundary corresponded to the critical temperature as defined by Wood and Knudsen, and the curvature of the boundary showed that this critical temperature decreases with decreasing stream density.

---

* 'Z. f. Physik,' vol. 25, p. 287 (1924).

† 'Z. f. Physik,' vol. 33, p. 340 (1925).

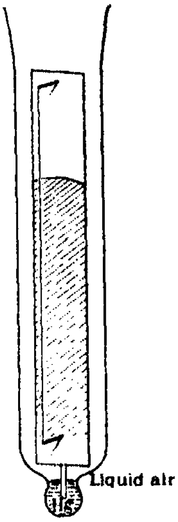


FIG. 1.

It was found, however, that for cadmium on mica this critical temperature could vary at least between —65° C. and —100° C. with different filament temperatures, whereas in the previous experiments it had been observed to vary only between —77° C. and —80° C. For silver streams produced in a similar manner the critical temperature varied between —90° C. and —15° C. It proved, however, to be impossible to obtain quantitative relations between the critical temperature and stream density by this method, since it was impossible either to maintain the filament at a constant temperature or measure its temperature accurately since the evaporation continually changes the resistance and temperature of the filament. A new apparatus was therefore constructed in which accurate measurements of the stream density were possible.

(3) *Experiments with Constant Surface Temperature.*

The principle of the new apparatus is to use a metallic oven as a source of vapour for the issuing stream. The vapour issues from a hole in the end of a metal tube so that its density has circular symmetry round the normal to the tube end surface, but decreases as the angle from the normal increases. The stream of varying density is allowed to fall on a condensing surface held at a constant and uniform temperature and the extent of the deposit observed. The boundary of the deposit fixes the critical stream density below which no film is built up for that particular surface temperature.

(i) *Apparatus.*—The apparatus is shown in fig. 2. The source of metallic vapour consisted of the small steel oven A. A piece of steel rod, approximately 8 mm. in diameter, was bored out to leave a very thin wall at the end, and here a fine circular hole H was bored by the finest available drill, giving a diameter of hole of 0·37 mm. The shape of the hole was examined under a microscope and found to be quite circular. When finer holes were required they were pricked by a fine needle in a sheet of thin steel foil and the foil spot welded to the end of the oven tube. In this way holes of less than 0·04 mm. diameter could be made. A screw G fixes an iron-constantin thermocouple in the steel wall to register the oven temperature and the end is closed by the screw B.

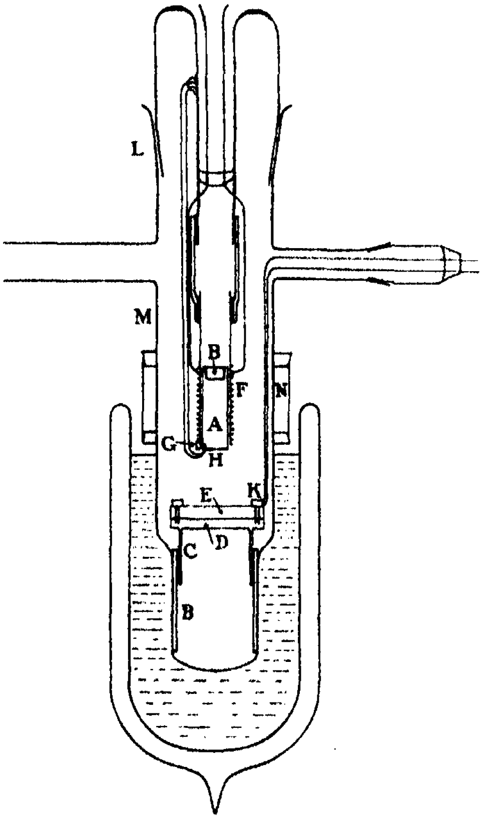


Fig. 2.

The heating element for the oven consists of a pyrex glass tube with a double row of projections holding in position a heating winding of nichrome ribbon. The pyrex tube is held centrally in the tube M by a ground glass joint L.

One of the most difficult points in the design of the apparatus was to find a method of keeping steady the temperature of the surface on which the atomic stream falls, this temperature having to be variable between the limits of −70° and −160° at least. It was finally solved by making a brass tube to fit as closely as possible to the constricted end of the tube M. Between the brass tube and the glass wall was placed a thin copper foil covered with Rose metal, a low melting point fusible alloy. The end of the tube was placed in a bath of heated oil until the Rose metal melted, and on cooling it then adheres closely to the glass and makes a very good thermal contact with it. In later apparatus better results were obtained by platinising the inner surface of the glass and copper plating. A copper tube C with a brass end plate D slides in the brass tube B. To the brass plate is screwed a copper or silver disc E, 2·5 cm. in diameter and 6 mm. thick. This plate is highly polished and serves as the collecting surface for the atomic stream from the oven. It also has a large thermal capacity and high thermal conductivity, so that the temperature in the plate is certainly uniform to within one degree when steady conditions are obtained.

The surface is cooled by immersing the whole of the outer tube in a flask of liquid air to the level of the oven mouth. This ensures that any cadmium atoms hitting the wall of the glass tube do not re-evaporate and interfere with the beam striking the plate. The actual steady temperature of the plate E is varied by varying the distance the copper tube C is inserted in the brass tube B. The temperature of the copper surface is measured by a copper-constantin thermocouple K. The whole apparatus is evacuated to a pressure of below 10^{-5} mm. of mercury by a mercury diffusion pump and Hyvac. A water jacket N keeps the walls of the glass tube, opposite the oven, cool.

(ii) *Experimental Procedure.*—The steel oven A is first cleaned and then outgassed at a temperature of 900° C. in a vacuum oven. A small quantity of cadmium is inserted, the screw B fixed and the oven placed inside the heater F. The copper plate E is polished by a succession of grades of emery paper until the finest available is reached. It is then washed with alcohol and any residual grease removed by vacuum ovening. It is now screwed to the plate D and inserted in the tube M by a special tool. The oven system is inserted and the distance from oven to copper plate adjusted, being usually between 1 and 2 cm. The apparatus is next evacuated and the oven heated to a temperature which is approximately that of the experiment and maintained at this temperature for about half an hour to ensure that any gas in the system is driven out. In the first series of experiments the receiving plate was not heated. The oven is then allowed to cool and the tube inserted in a flask of liquid air. The temperature of the copper plate falls quite slowly and it is usually about 1½ hours before a steady temperature is obtained.

The oven temperature is now raised slowly and the copper plate observed closely. When a certain oven temperature is reached, a small green circle appears on the copper plate opposite the hole in the oven if low stream densities are being used. The diameter of this circle grows and at the same time the colours in the central zones change, going through a complete sequence to red and finally attain a bluish metallic tint. If the oven temperature is allowed to increase, the whole of the copper plate finally gets covered with a metallic layer. A preliminary experiment of this type serves to fix approximately the density of the molecular stream at which a deposit begins to appear. The experiment is then repeated and the oven temperature brought to within 10 or 20 degrees of the temperature at which a deposit occurred in the preliminary experiment. It is maintained at this value for 10 to 20 minutes in the case of experiments with a high density of stream. If no deposit appears the oven temperature is then increased by 10 to 15 degrees, thus doubling the density of the stream. The circular deposit then usually appears and its growth is watched and measured with a travelling microscope. If the circle diameter is plotted against time a curve such as fig. 3 results, the circle diameter reaching an asymptotic value after a time varying from 1 or 2 minutes to several hours according to the stream density in use. Outside this circle no deposit at all occurs. Photograph 1, Plate 2, shows such a deposit.

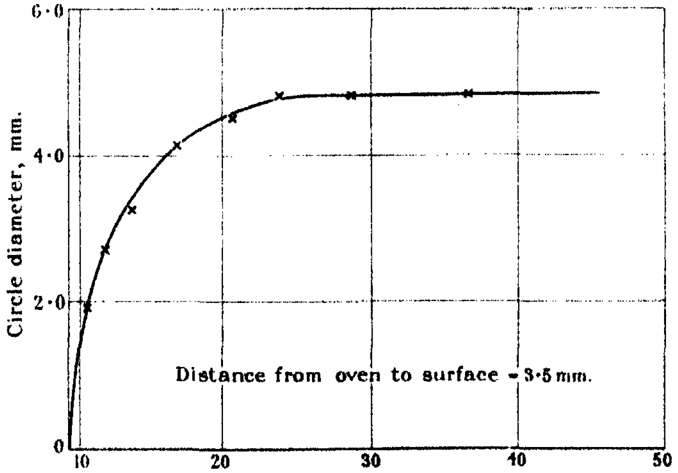


FIG. 3.

The reason for the circular deposit is evidently to be found in the cosine law of distribution of the atoms from the oven, the density of the stream striking the plate being a maximum at the centre and decreasing outwards with circular symmetry. The boundary of the deposit corresponds to a critical density below which no film can be built up, all points outside the circle having less than this critical density of incident stream.

If after the final circle diameter is reached the oven temperature is allowed to increase or the surface temperature to decrease, the growth of the circle begins again until a new steady condition is reached. If the stream density does not have circular symmetry in an experiment, then the forms of deposits obtained are not circular. Experiments were carried out, for example, in which a wire of 0·03 mm. diameter was stretched across the plate so that reflection from the wire occurred causing an increase of density on the surface to one side of the wire. The boundary of the deposit was found, as shown by photograph 2, to be moved outwards by about a millimetre where the wire crossed the boundary. A change of the boundary shape also occurs near the edge of the plate due to reflection of the stream from a hexagonal nut used to fix the plate in position.

(iii) *The Calculation of the Density.*—As is well known, the distribution of vapour emerging from a hole in a flat wall of a chamber is, under the conditions of molecular flow defined by Knudsen, given by a cosine law. The number of atoms emerging per second in a solid angle dw making an angle θ with the normal to the wall varies as $\cos\theta\, dw$.

The density of the atomic stream at a point B in the surface (fig. 4) is therefore

$$\nu = \mathrm{A}\cos\theta\,.\,\cos\theta/r^2 = \mathrm{A}\cos^4\theta/y^2.$$

The constant A is easily evaluated from gas kinetic theory and we find

$$\nu = \frac{p_0}{\sqrt{(2\pi m\mathrm{KT}_0)}}\cdot\frac{d\sigma\cos^4\theta}{\pi y^2}$$

where $d\sigma$ is the area of oven opening, p_0 and T_0 are the pressure and temperature inside the oven and m the mass of the cadmium atom. If then p is measured in millimetres of mercury and ν in atoms/sq. cm./sec.

$$\nu = 3\cdot3\,.\,10^{21}\frac{p\cos^4\theta\, d\sigma}{\pi y^2\sqrt{\mathrm{T}_0}}. \qquad (1)$$

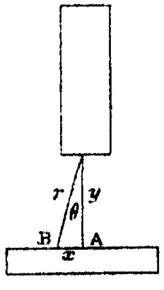


FIG. 4.

Knowing the diameter of the circular deposit, the distance y and the oven temperature, the density can therefore be calculated. The Knudsen condition for molecular flow necessary to make valid the relation (1) is that the mean free path of the vapour inside the oven should be greater than the dimensions of the oven opening. Care was therefore taken always to limit the oven temperature to such values that this condition was fulfilled. In most of the experiments the diameter of the hole in the oven was 0·4 mm., and the maximum oven temperature 260° C. giving a pressure of 10^{-2} mm. Hg and a free path of approximately 0·5 mm.

(iv) *The Relation between Critical Density and Temperature.*—Experiments have been carried out at a series of surface temperatures from —80° to —155°. In this region the critical density is found to vary from $2 . 10^{16}$ to $4 . 10^{12}$. At the lower temperature therefore it takes about 15 minutes to form a monomolecular layer, and since, according to Estermann,* the layer becomes visible only when, on the average, four atoms thick, it is about an hour before any sign of a deposit appears. To obtain a good metallic deposit and to ensure that the circle reaches its final diameter it is therefore necessary to carry on the experiment for several hours, keeping the oven temperature and surface temperature constant throughout this period. At the higher surface temperatures, on the other hand, a thick metallic deposit is built up in less than a minute.

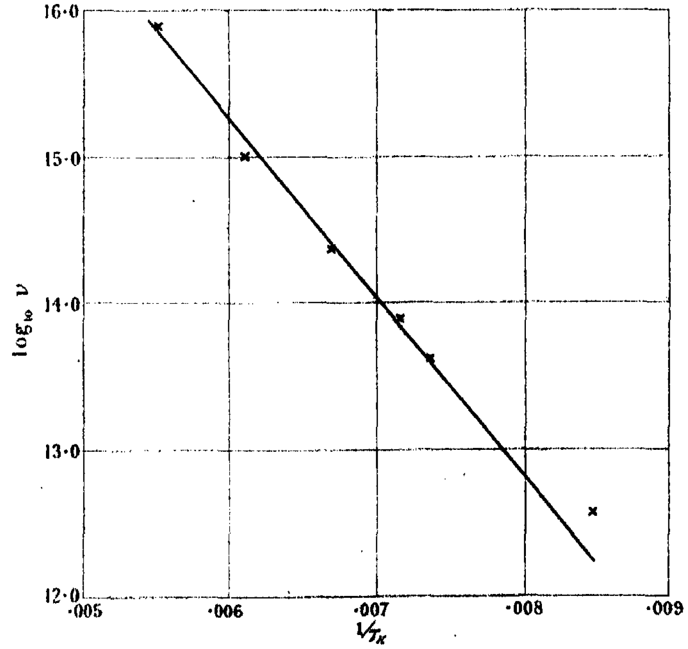


FIG. 5.

Mean values of the critical density observed are given in Table I.

Table I.

Surface temperature °C.	—92	—113	—126	—133	—137	—155
Critical density	$9 . 10^{15}$	$1·0 . 10^{15}$	$2·4 . 10^{14}$	$8 . 10^{13}$	$4 . 10^{13}$	$4 . 10^{12}$

---

* 'Z. Phys. Chem.,' vol. 106, p. 399 (1923).

With the rapid variation in density with surface temperature the numerical values cannot be repeated with any great accuracy but consecutive experiments generally agreed to within 20 per cent. at least.

In fig. 5 the logarithm of the density is plotted against the reciprocal of the absolute temperature of the surface, and it is seen that a practically linear relation holds, the relation being of the form

$$\nu = 4 \cdot 7 \, . \, 10^{22} \, e^{-2840/\mathrm{T}}. \tag{2}$$

(4) *Interpretation of Results.*

(i) *The Frenkel Theory of Adsorption.*—A theoretical investigation of the phenomena described by Chariton and Semenoff was made by Frenkel* which accounted for the existence of the critical temperature of condensation and its variation with density. He made two main assumptions: first, that the atoms on striking the surface stay on the surface for a definite time and during this time move about the surface like a two-dimensional gas; secondly, that when, during their motion on the surface, they collide with other atoms of the adsorbed vapour, the pair forms a nucleus with a much longer mean life on the surface than the single atoms, and these "doublet atoms" act as centres of condensation for other atoms. By an appeal to statistical mechanics he derived relations for the mean lives of the single and doublet atoms,

$$\tau = \tau_0 e^{u_0/\mathrm{KT}} \qquad \text{and} \qquad \begin{aligned} \tau_1 &= \tau' \, e^{(u_0 + \Delta u_1)/\mathrm{KT}} \\ &= \tau' \, e^{u_0/\mathrm{KT}} \end{aligned}$$

where u_0 is the energy of adsorption of a single atom on the surface and Δu_1 is the dissociation energy of a pair of atoms of the adsorbed gas, τ_0 and τ' being constants identified with the oscillation period of the adsorbed atom on the surface and T the temperature of the surface. If n be the number of adsorbed atoms per unit area of surface, σ the area within which the field of force of an adsorbed atom might be considered to be effective for capture, and σ_0 the gas kinetic theory molecular area, then under equilibrium condition the number of isolated atoms was assumed to be

$$n_0 = n\,(1 - n\sigma) \tag{3}$$

and the number of paired atoms

$$n_1 = n^2\sigma \qquad \text{where} \qquad \sigma = \sigma_0 \, e^{\Delta u_1/\mathrm{KT}} \tag{4}$$

---

* 'Z. f. Phys.,' vol. 26, p. 117 (1924).

these relations being derived from a mixture of arguments from thermodynamics and Kinetic Theory. If then ν be the number of atoms striking the surface per unit area in unit time, we have

$$\frac{dn}{dt} = \nu - \alpha n_0 - \alpha' n_1 = \nu - \alpha n + \beta n^2$$

where

$$\alpha = \frac{1}{\tau}, \qquad \alpha' = \frac{1}{\tau_1}$$

and

$$\beta = (\alpha - \alpha')\,\sigma_0\,e^{\Delta u_1/\mathrm{KT}}.$$

The condition for a stationary state is

$$\frac{dn}{dt} = 0,$$

and since (4) may be written

$$\frac{dn}{dt} = \beta\,(n - \alpha/2\beta)^2 + (\nu - \alpha^2/4\beta) \tag{6}$$

if $\nu < \alpha^2/4\beta$ then clearly as n increases to $\alpha/2\beta$, dn/dt becomes zero and would be negative for larger n. It is therefore impossible for a layer to be built up.

If $\nu > \alpha^2/4\beta$, dn/dt is essentially positive and a deposit will be formed. The condition for the "critical density" is therefore given by

$$\nu_c = \alpha^2/4\beta = \frac{1}{4\sigma_0\,\tau_0\,e^{u_1/\mathrm{KT}}},\quad u_1 = u_0 + \Delta u_1 \tag{7}$$

since $\alpha'/\alpha = e^{-\Delta u_1/\mathrm{KT}}$ will be small.

We see that this relation is of the form found experimentally and given by (2). We have therefore verified that the relation between critical temperature and density predicted by the Frenkel theory is true within the limits of experimental error between −90° C. and −150° C.

We may proceed further and find the numerical values of the constants u_1 and τ_0 from our results. By identifying (2) and (7) we have immediately

$$u_1/\mathrm{K} = 2840$$

so that $u_1 = 5680$ calories/mol.

This energy represents the energy required to remove one member of a cadmium "doublet" from the surface. It is the sum of u_0, the energy of an isolated atom on the surface and the "heat of dissociation" of the doublet. The point of greatest interest would be to separate u_0 and Δu_1, for there is no other experimental evidence on the heat of dissociation of a cadmium doublet or molecule, but from the data given by the critical densities and temperatures alone, we cannot find this and must look for other means of separating the factors.

We have also $4\sigma_0\tau_0 = 1/4{\cdot}7\,.\,10^{22}$ and taking $\sigma_0 = 2\,.\,10^{-15}$, $\tau_0 = 2{\cdot}6\,.\,10^{-9}$ secs. This gives for the mean life of a paired atom in the surface at a temperature of $-100°$ C.

$$\tau_1 = 2{\cdot}6\,.\,10^{-9}\,e^{2840/173} = 3{\cdot}4\,.\,10^{-2}\text{ secs.}$$

The mean life of an isolated atom cannot be found until we can split up u_1 into its components u_0 and Δu_1. In view of the lack of rigidity of the arguments by which the relations (4) were derived however, little reliance can be attached to the numerical values of τ_1 and τ_0 derived.

[*Added May* 1, 1928.—We should further point out that the verification of the form (7) for ν_c is not necessarily a proof of the assumption that the centres of higher stability are doublets. If, for example, we were to assume that these centres are only formed by groups of seven atoms, then in the equation for dn/dt we should replace βn^2 by $\beta_1 n^7$, and it can be shown that the critical density would again be given by an expression similar to (7) but with a different numerical constant. We should expect however that these centres of greater complexity would have a value for Δu approaching 32,000 calories/mol., the heat of sublimation of a cadmium crystal, and the small value for u_1 actually obtained, rather supports the view put forward by Frenkel.]

(iii) *Relation between the Time of Appearance of a Deposit and the Stream Density.*—If we can, on the basis of the Frenkel Theory, account satisfactorily for the curves of growth of the circles observed and exemplified by fig. 3 we should expect to be able to use the theory to determine the proportion of atoms which re-evaporate during the building up of a visible deposit.

The time taken to build up a deposit of just visible intensity was observed in different experiments by first allowing a circular deposit to grow and then by means of the ground joint moving the oven position and observing the time for the first appearance of the centre of a deposit under the new oven position. Experiments of Estermann (*loc. cit.*) have shown that a deposit is first visible when it is on the average about four atoms thick. If this surface density be denoted by n_0 then the time of appearance t_0 should be n_0/ν. This was actually observed to hold over a very wide range of stream densities.

If this relation held for all densities, however, the time taken to form a deposit at a point in the surface fixed by an angle θ would be $t_0/\cos^4\theta$, and we should observe a curve of growth similar to curve A of fig. 6. Actually, however, as shown by fig. 3 a much longer time is taken to build up a just visible deposit as the density approaches the critical density.

To calculate the time of appearance from the Frenkel Theory outlined in section (i) we must first find the time to reach a surface density $n_c = \alpha/2\beta = 1/\sigma$

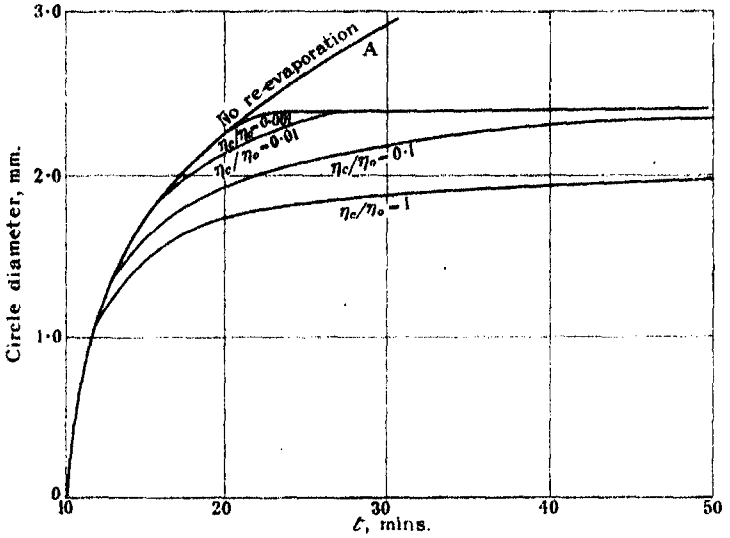


FIG. 6.

since above this surface density the number of isolated atoms must be zero, $n_c\sigma$ being unity, so that every atom striking the surface will come within the radius of attraction of another adsorbed atom and will therefore have a much longer mean life.

The time to reach the density n_c will be from (4)

$$t_1 = \int_0^{\alpha/2\beta} \frac{dn}{\nu - \alpha n + \beta n^2} = \frac{1}{\sqrt{(\beta\nu - \alpha^2/4)}} \text{ arc tan } \frac{1}{\sqrt{(4\beta\nu/\alpha^2 - 1)}}$$

or writing $\nu = \lambda\nu_c$ where ν_c is the critical density $\alpha^2/4\beta$

$$t_1 = \frac{2}{\alpha\sqrt{(\lambda - 1)}} \cdot \text{arc tan } \frac{1}{\sqrt{(\lambda - 1)}} = \frac{n_c\lambda}{\sqrt{/(\lambda - 1)}} \text{ arc tan } \frac{1}{\sqrt{(\lambda - 1)}}.$$

For λ very large or for densities large compared with the critical, $t_1 = n_c/\nu$, *i.e.*, there is practically no re-evaporation. The time taken for the surface density to increase from n_c to n_0 will be $(n_0 - n_c)/\nu$ since during this time re-evaporation will be small and the total time to form the surface density $n_0 = (t_1 + t_2)$. The main point of interest is to know $(t_1 + t_2)/t_0$, the ratio of the actual time taken for a deposit to appear to the time if no re-evaporation occurred. This ratio is tabulated as a function of ν/ν_c and n_0/n_c in Table II.

Table II.—$(t_1 + t_2)/t_0$.

n_0/n_c.	$\lambda = \nu/\nu_c$.	100	10	4	1·5	1·1	1·01	1·001
1		1·005	1·07	1·21	6·5	17·1	58·4	198
10		1·000	1·007	1.021	1·55	2·61	6·74	20·7
100		1·000	1·001	1·002	1·055	1·16	1·57	2·97

We can now account for the curves of growth observed. As we move along a radius from the centre of a deposit the density of the stream approaches more and more closely to the critical density so that from Table II we can see that the time taken to build up a visible layer increases rapidly. Fig. 6 gives the actual form of curve which would be obtained from Table II for different values of n_0/n_c and we see that a curve similar to that of fig. 3 is obtained if n_0/n_c lies between 100 and 1000. Since $n_0/n_c = 2n_0\beta/\alpha = 2n_0\sigma_0\, e^{\Delta u_1/\mathrm{KT}}$ and since for a just visible layer $n_0\sigma = 3$, this would correspond to a value of Δu_1 of about 2000 calories per mol., but in view of the roughness of the method and the lack of rigidity in the arguments on which the basic relation (4) was founded no great value is to be attached to this determination.

The point of definite practical interest which emerges is that the Frenkel Theory gives a sufficiently correct prediction of the observed phenomenon for its results as expressed in Table II to be used for the determination of the proportion of atoms re-evaporating during the formation of a film. Since we may take it that n_0/n_c will always be greater than 10 it follows that provided the density of the atomic stream is at least four times the critical density for the surface temperature used the proportion of atoms re-evaporating may be neglected.

(4) *The Effect of the Surface on the Critical Temperature and Density.*

(i) *The Critical Density for Different Surface Materials.*—The effect of a change in the surface was investigated by soldering to half of one of the copper discs a thin plate of silver. In this way the surface temperature and stream density were exactly the same for the two metals. Photograph 3 shows the deposit formed at a temperature of —137° C. The circular boundary of the deposit is practically continuous at the junction of the silver and copper surfaces.

We have seen from photograph 2 that the slight change in density resulting from the reflection of the stream from a 1/30th mm. wire is sufficient to cause an appreciable change in the shape of the boundary so that it is evident without calculation that the critical densities are the same for silver and copper at this surface temperature to within 1 or 2 per cent. The experiments were repeated at a temperature of −170° C. and exactly the same continuity of the circular boundaries was found.

A thin cover glass was then attached to one half of the upper plate and a deposit formed. The result is shown in photograph 4. It is evident that in this case also the critical densities for the two surfaces are practically identical. The thickness of the cover glass was 0·1 mm. so that there could be no appreciable temperature difference between the copper and glass surfaces. These photographs show in a most direct manner that under the conditions of the experiment the nature of the surface affects the phenomena very little.

We must therefore conclude that the values of u_1 for the three surfaces are practically identical. It is hardly to be expected that the values of the attraction of the different surfaces for a cadmium atom are identical so that we can only assume either that u_0 is small compared with Δu_1 in the three cases, or that the true surface forces are masked by some contaminating influence such as a gas film. Since the object of the first experiments was to obtain results under the conditions usually prevailing in experiments with molecular streams, no attempt was made to outgas the surface, but it became evident that further work was necessary to determine the effect of different surface conditions on the phenomena.

(ii) *Experiments to show the Effect of an Uncontaminated Surface.*—An apparatus was now constructed by which a new silver surface could be deposited immediately before or during the formation of the cadmium deposit. A quartz tube wound with a heating spiral of platinum strip and containing a small pellet of silver was fixed by the side of the cadmium oven so that a new silver layer could be deposited at any time. With a steady receiving surface temperature, the cadmium oven was heated to a temperature giving a stream density below the critical density for the surface. It was maintained at this temperature for a sufficient time to verify that no deposition was occurring. The silver oven was now heated and simultaneously with the formation of a faint circular deposit under the silver oven indicating that the deposition of metallic silver had begun, a deposit of cadmium began to form and spread rapidly over that part of the surface exposed to the silver beam, the centre of the circular cadmium deposit being under the nozzle of the silver oven. The obvious interpretation of these experiments is that the newly deposited silver remains uncontaminated for a sufficient time for it to serve as a nucleus for a cadmium deposit to grow, the cadmium atom having a much longer mean life on the uncontaminated surface than on the original silver surface.

In further experiments ground joints were eliminated from the apparatus, and the junction between the glass and the tube B of fig. 2 made by platinising and copper plating the interior of the glass tube, so that before each experiment the whole apparatus could be baked out for 2 hours at 350° C. In addition to this, the new silver surface was deposited as in the previous experiments, the silver pellet being now contained in a steel tube placed inside the heated quartz tube, this arrangement giving better dispersion of the silver. It was now possible to be certain that the new silver surface would not be contaminated by adsorption of gas, and it was found that a decrease by a factor of 10 occurred in the density required to form a deposit.

Sufficient experiments have not been carried out to determine what the value of u_1 will be under these surface conditions, since very long exposures will probably be necessary at the lower temperatures to determine whether or not a critical density exists.

The purer the surface conditions the greater might be expected to be the value of u_1. The heat of sublimation of a cadmium crystal is of the order of 32,000 calories per mol., and we might expect that u_1 would tend to a figure of this order for cadmium on an absolutely pure silver or copper surface.

(iii) *Other Experiments.*—Further support for the hypothesis that the true surface forces are masked is given by the fact that Estermann (*loc. cit.*) obtained widely different figures for u_1 in an investigation on the critical densities for cadmium atoms on copper, glass and silver at three temperatures between −83° and −111° C.

Cadmium was evaporated from an oven and after passing through two collimating slots fell on a cooled receiving surface. The oven temperature was maintained constant at a given figure and the surface temperature lowered in steps of between 5 and 10°. If at any one surface temperature a deposit did not appear in about 5 minutes it was assumed that the surface temperature was above the critical temperature, and it was lowered and a further wait of 5 to 10 minutes made. The first appearance of a deposit fixed the critical temperature and density.

The following figures were obtained for cadmium on copper :—

T_c	−111° C.	−93° C.	−88° C.
ν_c	$3\cdot4 . 10^{13}$	$7\cdot7 . 10^{13}$	$12 . 10^{13}$

giving

u_1 = 2910 cal./mol. from the first and second points,
3900 cal./mol. from the second and third points.

The density figures for cadmium on silver were about one-quarter of those on copper, and for cadmium on glass were about 10 per cent higher than for cadmium on copper.

It is seen that much smaller values for u_1 are obtained, and although the method of Estermann was not such as to allow of accurate density determinations the difference is too great to be explained by experimental error. Estermann's surface conditions differed from those in our experiments in that the condensing surface was preheated to 250° C. before the experiment. The use of ground glass joints, however, made it impossible to bake out the apparatus, and it is extremely unlikely that gas films were eliminated.

Other evidence showing the effect of gas films was given by Langmuir* in an investigation on the condensation of cadmium vapour on glass. Cadmium condensation nuclei were first formed on the walls of a bulb containing cadmium vapour by applying liquid air at one point. The liquid air was removed, and it was found that when the pressure of the residual gas was reduced by applying liquid air to a charcoal tube, a very uniform deposit was formed at the place where the nuclei existed. When the liquid air was taken off the charcoal however the deposit formed in a much less regular manner.

(5) *Experiments to Show the Reality of the Surface Motion.*

One of the fundamental assumptions of Frenkel was that the atoms, after striking the surface, do not remain in one fixed position but move about and are able to form doublets. This view is supported by experiments carried out by Estermann (*loc. cit.*) on the deposits formed when V-shaped diaphragms are placed in the track of the molecular beam.

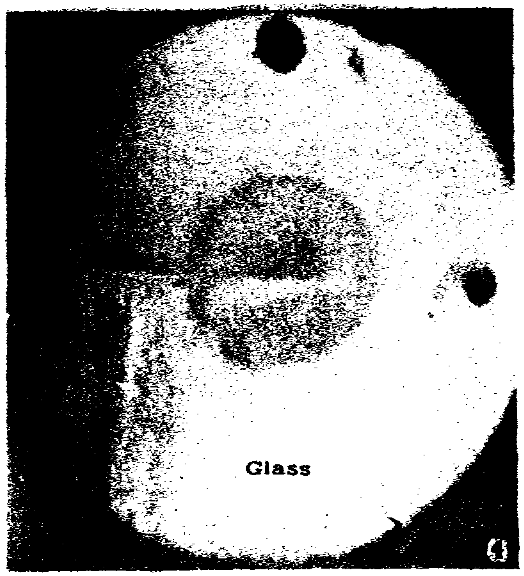
It was found that no deposit appeared corresponding to the narrow part of the diaphragm when the surface temperature was just below the critical temperature for the stream density used. In other words, the critical density is higher at the edge of the deposit than at some point in the centre. This is most easily explained by assuming that lateral motions of the atoms occur after striking the surface and whilst at points in the centre of the deposit as many atoms move into an area as leave it, at the edge no replacement occurs and so the effective density is lowered.

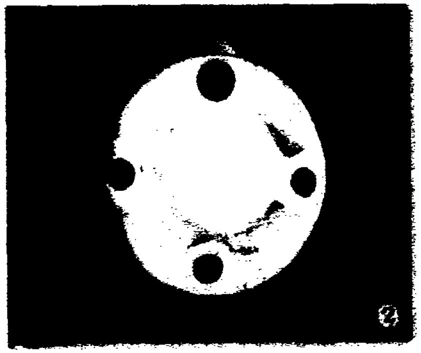
Further support is given to the hypothesis by the work of Volmer† on the great differences between the rate of growth of crystals along different planes,

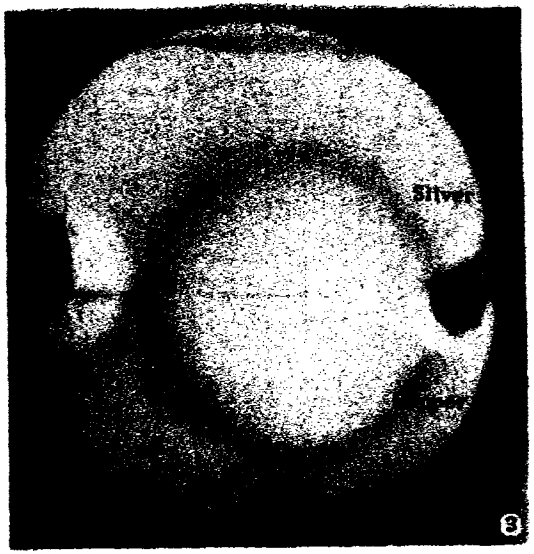
---

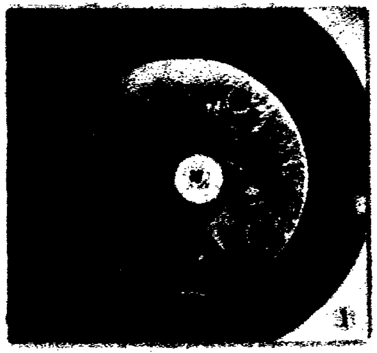
* 'Proc. Nat. Acad. Sci.,' vol. 3, p. 141 (1917).

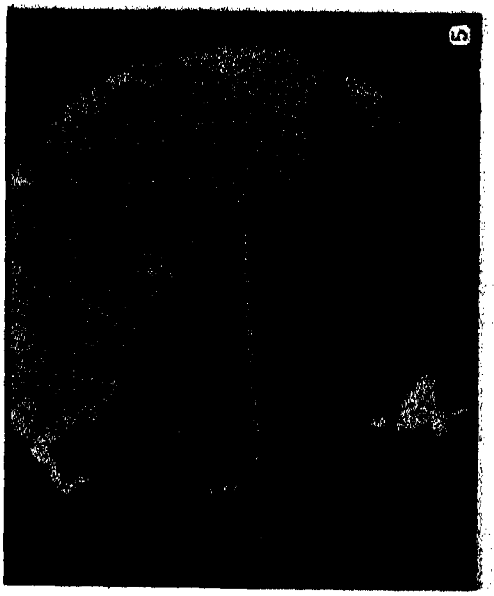
† 'Z. f. Physik,' vol. 9, p. 193 (1922).


Glass
1


2


Silver
3


4


5

6

7

8

this difference being most satisfactorily explained by migrations of the adsorbed atoms from the plane of incidence to another plane.

An attempt was made by the author to get a more direct verification of the reality of surface motion. To this end a 0·015 mm. wire was placed as close to the surface as possible, in the path of the atomic beam, the arrangement being sketched in fig. 7.

An oven was used with a very fine opening to get a well defined shadow region SS. The deposit was formed in the usual way. If then no motion of the atoms along the surface occurs, we should get a region SS where no deposit at all is visible. In the first experiments no deposit was actually found in the shadow zone. It was thought, however, that this was no disproof of the creeping hypothesis since it is quite likely that the density of the atoms in the region SS would be below the critical density necessary for a layer to be built up.

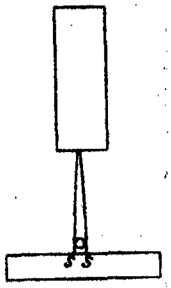


FIG. 7.

A method of overcoming this difficulty is to "sensitise" the surface before putting the wire in position. Langmuir (*loc. cit.*) showed that if a metallic vapour is allowed to condense on a cooled surface for a short time, so as to leave condensation nuclei without any visible film being apparent, and if later the surface temperature is raised and a vapour stream allowed to fall on it, then condensation will occur on the nuclei and on no region other than that "sensitised" in this way.

The method of sensitising adopted here was to place the wire in position and begin the formation of a circular deposit. As soon as a circle of diameter AA′ say (fig. 8) became visible, the wire was moved to a different position ZX on the plate, by an electromagnet. The growth of the circle then continued until the limiting diameter BB′ was reached. When examined under a high power microscope it was now found that crystals of cadmium had accumulated under the shadow of the wire in the position XZ between C and C′. It is evident that when the circle had reached a diameter AA′, a certain zone outside this circle would have a layer of condensation nuclei, or would, in other words, be sensitised although no deposit was visible. On moving the wire to the position XZ, therefore, the shadow zone would possess condensation nuclei to collect any

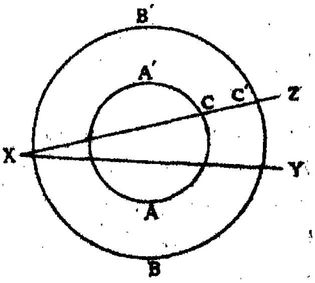


FIG. 8.

atoms creeping into the region and we might expect a deposit to be formed if any creeping exists. Microphotographs of the deposit found under the wire are shown in photographs 5, 6, and 7 (Plate 3).

Photograph 5 is a low power picture of the edge of the final circle BB′ near the wire shadow. The wire has been moved from a position Y to the position shown by the white shadow CDE. Photographs 6 and 7 are taken with the high power at points C and D. Crystals of cadmium are very definitely seen in the shadow zone. The diameter of the circle AA′ before the wire was moved was 4·6 mm. and the final diameter AA′ 9·7 mm., so that there is no possibility of the deposit under the wire at C and D being there before the wire was moved.

The results therefore strongly support the hypothesis of Frenkel.

Incidentally, microphotograph 5 shows very clearly the effect of sensitising on a surface. It will be seen that at the point Y on either side of the *original* position of the wire the boundary of the deposit has moved outwards. The only explanation for this fact is that the reflection of the stream from the wire increases the density on either side of the wire and leaves a sensitised region of greater diameter along these two lines so that when the wire is moved, more deposition occurs along these lines than on the rest of the surface.

Further indirect evidence in favour of surface motion is obtained from the optical properties of the films formed. During the experiments with streams of low density it is found that the deposit assumes very brilliant colours during the early periods of condensation. Usually the first circle seen is green and later turns to red. In some cases a double sequence of colours is observed on the final film, the colours in one case being blue, red, yellow, green, blue, etc., in concentric zones. After a succession of colours the central zone turns to a bluish metallic tint with cadmium, and the colours disappear.

When stream densities of the order of 10^{15} atoms/sq. cm. per second are used no colours are seen except at the edge of the final circle.

In experiments in which a stream density of 10^{13} atoms per square centimetre per second was used, the coloured circles appeared in about 15 minutes. Now to form a monomolecular layer requires about 3.10^{15} atoms per square centimetre, so that it would take 5 minutes to form such a layer with the stream density used. It is well known that a layer of three atoms thick is just visible with many metals, so that our experiments agree in assigning an average thickness to these coloured films of about three atoms. It is therefore quite impossible that the colours are due to interference. Volmer and Estermann* showed, by ultra-microscopic methods, that these very thin films consisted not of a uniformly distributed layer, but of a number of small isolated crystals, one crystal containing up to a thousand atoms in some cases.

---

* 'Z. f. Physik,' vol. 7, p. 1 (1921).

In our experiments, therefore, the colours are the "resonance colours" described by Maxwell Garnett* and others, being properties of the individual crystal size and the surface density of the crystals. The mechanism of the formation of a surface layer consisting of these isolated crystals can only be provided by the theory of the surface motion of the atoms in the adsorbed state, an isolated atom moving about the surface until it collides either with another atom to form a pair of much longer life or with a nucleus of higher complexity and still greater stability.

Photograph 8 shows one of the most striking systems of rings observed. One ring is quite isolated from the remainder of the deposit, and its position does not seem to change with the type of light used for viewing the surface so that it is not merely an optical effect. Whether it is due to a real maximum in the density due to selective reflection from the walls of the oven opening where cadmium crystals accumulate is not yet certain, but further experiments are in progress. The wave length to be associated with the cadmium atom at 300° C. on the Wave Mechanics is about 10^{-9} cms. so that it is possible that such selective reflection might be explained in this manner.

(6) *The Effects of External Fields on the Phenomenon.*

During the experiments with silver streams described in Section 2, an attempt was made to see whether the application of an intense magnetic field to the surface would change the surface forces sufficiently to produce a change in the form of the boundary of the deposit, as it was thought possible that orientation of the silver atom in the field might be sufficient to produce such a change. Pole pieces were inserted inside the evacuated tube and a field of the order of 10,000 gauss produced at one point of the boundary, by a sharp edged pole piece placed behind the condensing plate. No change in the boundary shape was observed however. It might be thought that an external electric field would be a more promising source of interest, but since the surface fields are probably of the order of 10^8 volts/cm. it is impossible, by ordinary means, to produce fields likely to alter appreciably the surface forces. It would, however, be interesting to observe the phenomenon on the surfaces of the polarised crystals of rock salt or quartz described by Joffé† where polarisation fields of this order of magnitude obtain.

---

* 'Phil. Trans.,' vol. 203, A, p. 385 (1904).

† 'Ann. d. Physik,' vol. 72, p. 461 (1923).

(7) *Conclusions.*

The experiments described have shown that most of the phenomena occurring in the building up of films by condensation of atomic streams can be explained by the general picture of the adsorption process put forward by Frenkel. If this theory is accepted then the method described enables the magnitude of the surface forces to be found for different adsorbed atoms and different surface materials. It has been shown that the true surface forces are in general masked by contaminating influences which are almost certainly gas films, and that if the magnitude of the real surface forces is required then special precautions have to be taken to eliminate gas from the apparatus and to produce a new surface before beginning the condensation. If these precautions are taken the method should prove very fruitful in the investigation of the adsorption forces for a whole series of vapours and surfaces.

The effect of gas films has been shown to influence very considerably the critical temperature phenomena described by Wood and Knudsen, and it is probable that if these films are eliminated the phenomena would not exist for many surfaces. In most experiments with molecular streams these films are not eliminated, and the experiments described show that to ensure that practically all the atoms of the stream condense it is sufficient to have the density of the stream of the order of four times the critical density for the surface, but that provided liquid air temperatures are used for the condensing surface the critical atomic stream density for cadmium may be so low as to require more than 10 hours for the first appearance of a deposit.

The extreme sensitivity of the boundary of the deposits to changes in stream density may be a valuable tool in experiments on the reflection of atomic streams from crystal and other surfaces where it is desirable to obtain information on the symmetry of the reflected beam, and it is hoped to use this property in experiments now in progress.

I have finally to express my thanks to Dr. P. Kapitza for many valuable suggestions and to Sir Ernest Rutherford for his continued interest in the work.

The Stark Effect of the Fine-Structure of Hydrogen.

By R. Schlapp, M.A., Ph.D., Lecturer in Applied Mathematics, University of Edinburgh.

(Communicated by C. G. Darwin, F.R.S.—Received April 2, 1928.)

One of the earliest successes of classical quantum dynamics in a field where ordinary methods had proved inadequate was the solution, by Schwarzschild and Epstein, of the problem of the hydrogen atom in an electric field. It was shown by them that under the influence of the electric field each of the energy levels in which the unperturbed atom can exist on Bohr's original theory breaks up into a number of equidistant levels whose separation is proportional to the strength of the field. Consequently, each of the Balmer lines splits into a number of components with separations which are integral multiples of the smallest separation. The substitution of the dynamics of special relativity for classical dynamics in the problem of the unperturbed hydrogen atom led Sommerfeld to his well-known theory of the fine-structure of the levels; thus, in the absence of external fields, the state $n = 1$ ($n = 2$ in the old notation) is found to consist of two levels very close together, and $n = 2$ of three, so that the line H_α of the Balmer series, which arises from a transition between these states, has six fine-structure components, of which three, however, are found to have zero intensity.

The theory of the Stark effect given by Schwarzschild and Epstein is adequate provided that the electric separation is so much larger than the fine-structure separation of the unperturbed levels that the latter may be regarded as single; but in weak fields, when this is no longer so, a supplementary investigation becomes necessary. This was carried out by Kramers,* who showed, on the basis of Sommerfeld's original fine-structure theory, that the first effect of a weak electric field is to split each fine-structure level into several, the separation being in all cases proportional to the square of the field so long as this is small. When the field is so large that the fine-structure is negligible in comparison with the electric separation, the latter becomes proportional to the first power of the field, in agreement with Schwarzschild and Epstein. The behaviour of a line arising from a transition between two quantum states will be similar;

---

* 'Z. f. Physik,' vol. 3, p. 199 (1920).

each of the fine-structure components will first be split into several, with a separation proportional to the square of the field; as the field increases the separations increase, and the components begin to perturb each other in a way which leads ultimately to the ordinary Stark effect.

The advent of wave-mechanics made further work necessary. A theory of the Stark effect, comparable with that of Schwarzschild and Epstein in that it ignored the fine-structure, was given by Schrödinger,* using wave mechanics. As was to be expected, the results were in substantial agreement with those of the older theory. It now appears desirable, since Sommerfeld's original explanation of the fine-structure has been superseded, to reconsider from the new point of view the problem dealt with by Kramers.

The purpose of this paper is to investigate the interaction between fine-structure and Stark effect by a method embodying the new theory of fine-structure which has developed out of the idea of the "spinning electron." The work is based on the wave equations recently given by Darwin† and subsequently given a final form by Dirac.‡ The characteristic feature of these equations is that they consist of two second order equations, or four first order equations; it is in this way that the duality which is necessary to explain the observed spectra is introduced.† The equations were shown to account for fine-structure and the Zeeman effect. In applying them here to the corresponding problem of the Stark effect, we shall confine our attention chiefly to the two extreme cases of weak and strong fields, although, of course, the method is perfectly general. We shall, in other words, examine the Stark effect of the fine-structure components in weak fields, and the fine-structure of the Stark components in strong fields. We may point out at the outset that throughout this paper the terms "weak" and "strong" as applied to fields will be used in a relative sense. A field will be called weak or strong according as the separation it produces in levels or lines originally coincident is small or large compared with the fine-structure separation. A field which is strong in this sense may, as a matter of fact, still be, and in the "ordinary" as opposed to the "quadratic" Stark effect always is, very much smaller than the nuclear field acting on the electron.

§ 2. To a sufficient approximation the equations satisfied by the two wave-functions f and g of the electron under an electric intensity $(E_xE_yE_z)$ are, in the absence of a magnetic field,

---

* 'Ann. d. Physik,' vol. 80, p. 437 (1926).

† 'Roy. Soc. Proc.,' A, vol. 116, p. 227 (1927).

‡ 'Roy. Soc. Proc.,' A, vol. 117, p. 610 (1928).

$$\mathrm{D}f + \tfrac{1}{2}\frac{e}{mc^2}\Big[-i\Big(\mathrm{E}_y\frac{\partial}{\partial z} - \mathrm{E}_z\frac{\partial}{\partial y}\Big)g - \Big(\mathrm{E}_z\frac{\partial}{\partial x} - \mathrm{E}_x\frac{\partial}{\partial z}\Big)g$$

$$+ i\Big(\mathrm{E}_x\frac{\partial}{\partial y} - \mathrm{E}_y\frac{\partial}{\partial x}\Big)f + \Big(\mathrm{E}_x\frac{\partial}{\partial x} + \mathrm{E}_y\frac{\partial}{\partial y} + \mathrm{E}_z\frac{\partial}{\partial z}\Big)f\Big] = 0,$$

$$\mathrm{D}g + \tfrac{1}{2}\frac{e}{mc^2}\Big[-i\Big(\mathrm{E}_y\frac{\partial}{\partial z} - \mathrm{E}_z\frac{\partial}{\partial y}\Big)f + \Big(\mathrm{E}_z\frac{\partial}{\partial x} - \mathrm{E}_x\frac{\partial}{\partial z}\Big)f$$

$$- i\Big(\mathrm{E}_x\frac{\partial}{\partial y} - \mathrm{E}_y\frac{\partial}{\partial x}\Big)g + \Big(\mathrm{E}_x\frac{\partial}{\partial x} + \mathrm{E}_y\frac{\partial}{\partial y} + \mathrm{E}_x\frac{\partial}{\partial z}\Big)g\Big] = 0,$$

where

$$\mathrm{D} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Big(\frac{2\pi mc}{h}\Big)^2 + \Big(\frac{2\pi e\mathrm{V}}{ch}\Big)^2 - 2\,.\,\frac{2\pi i}{h}\,.\,\frac{e\mathrm{V}}{c^2}\frac{\partial}{\partial t},$$

V being the electrostatic potential, and N the atomic number. In the present problem we have the external electric field F superposed on the nuclear field, so that

$$\mathrm{V} = \frac{\mathrm{N}e}{r} - \mathrm{F}z \text{ and } \mathrm{E}_x = -\frac{\partial \mathrm{V}}{\partial x} = \frac{\mathrm{N}e}{r^2}\,.\,\frac{x}{r},\quad \mathrm{E}_y = \frac{\mathrm{N}e}{r^2}\,.\,\frac{y}{r},\quad \mathrm{E}_z = \frac{\mathrm{N}e}{r^2}\,.\,\frac{z}{r} + \mathrm{F}.$$

Note that e is taken positive, so that the charge of the electron is $-e$. These are the same as Darwin's original equations, except for the terms in

$$\mathrm{E}_x\frac{\partial}{\partial x} + \mathrm{E}_y\frac{\partial}{\partial y} + \mathrm{E}_z\frac{\partial}{\partial z},$$

which are new.†

The external field being much smaller than the nuclear field on the electron, it is at once apparent that the most important occurrence of F is in the term $-2\,.\,\frac{2\pi i}{h}\,.\,\frac{e\mathrm{V}}{c^2}\frac{\partial}{\partial t}$ in D, roughly speaking because $\frac{\partial}{\partial t}$ brings in a factor c^2. The equations therefore become, to this degree of approximation,

$$\mathrm{D}f + \tfrac{1}{2}\frac{e}{mc^2}\,.\,\frac{\mathrm{N}e}{r^3}[-i\mathrm{R}_1 g - \mathrm{R}_2 g + i\mathrm{R}_3 f + \mathrm{R}_4 f] = 0,$$

$$\mathrm{D}g + \tfrac{1}{2}\frac{e}{mc^2}\,.\,\frac{\mathrm{N}e}{r^3}[-i\mathrm{R}_1 f + \mathrm{R}_2 f - i\mathrm{R}_3 g + \mathrm{R}_4 g] = 0,$$

where

$$\mathrm{D} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Big(\frac{2\pi mc}{h}\Big)^2 + 2\,.\,\frac{2\pi i}{c^2 h}\Big(\frac{\mathrm{N}e^2}{r} - e\mathrm{F}z\Big)\frac{\partial}{\partial t} + \Big(\frac{2\pi \mathrm{N}e^2}{ch}\Big)^2\frac{1}{r^2},$$

$\mathrm{R}_1\mathrm{R}_2\mathrm{R}_3$ are the xyz components of the operator $\Big[r\,.\,\frac{\partial}{\partial r}\Big]$ and $\mathrm{R}_4 = \Big(r\frac{\partial}{\partial r}\Big)$.

---

† It is perhaps worth mentioning that all the results of this paper were worked out using Darwin's original equations (*loc. cit.*) which fail only when $k = 0$; it was, however, easy to see in the present instance how to amend them so as to give the correct levels. Dirac's equations, even in the approximation to which we use them here, hold for all values of k; the additional terms which they contain are, as we shall see, without effect unless $k = 0$.

§ 3. The process of solution is similar to that in Darwin's paper.† Write

$$f = \sum_{k,u} a_{k,u} \psi_{k,u}, \qquad g = \sum_{k,u} b_{k,u} \psi_{k,u},$$

where

$$\psi_{k,u} = f_n{}^k\left(\frac{4\pi^2\mathrm{N}me^2}{h^2(n+1)}r\right)\mathrm{P}_k{}^u(\cos\theta)\exp(iu\phi)\exp\left\{-\frac{2\pi i}{h}\left(mc^2-\frac{2\pi^2\mathrm{N}^2me^4}{h^2(n+1)^2}+\overline{\mathrm{W}}\right)t\right\}.$$

We may quote the definitions of the functions $f_n{}^k$ and $\mathrm{P}_k{}^u$ for reference,

$$f_n{}^k(z) = e^z z^k\left(\frac{d}{dz}\right)^{n+k+1}\frac{e^{-2z}z^{n-k}}{(-2)^{n+k+1}} \qquad (k \leqslant n),$$

$$\mathrm{P}_k{}^u(z) = (k-u)!\,(1-z^2)^{\frac{u}{2}}\left(\frac{d}{dz}\right)^{k+u}\frac{(z^2-1)^k}{2^k\,k!} \qquad (-k \leqslant u \leqslant k).$$

Carrying out the substitution, we find, to the required degree of approximation,

$$\frac{\mathrm{D}\psi_{k,u}}{\psi_{k,u}} = \frac{8\pi^2m}{h^2}\{\overline{\mathrm{W}}+\mathrm{C}(r)-e\mathrm{F}r\cos\theta\},$$

where

$$\mathrm{C}(r) = \frac{8\pi^4\mathrm{N}^4e^8m}{c^2h^4(n+1)^4}\left\{\frac{1}{4}-\frac{(n+1)^2h^2}{4\pi^2\mathrm{N}e^2m}\frac{1}{r}+\left(\frac{(n+1)^2h^2}{4\pi^2\mathrm{N}e^2m}\right)^2\frac{1}{r^2}\right\}$$

Hence our equations become

$$\sum\left[\frac{8\pi^2m}{h^2}\{\overline{\mathrm{W}}+\mathrm{C}(r)-e\mathrm{F}r\cos\theta\}-\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{u}{r^3}+\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{1}{r^2}\frac{\partial}{\partial r}\right]a_{k,u}\psi_{k,u} \\ -\sum\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{k+u}{r^3}b_{k,u}\psi_{k,u-1} = 0,$$

$$\sum\left[\frac{8\pi^2m}{h^2}\{\overline{\mathrm{W}}+\mathrm{C}(r)-e\mathrm{F}r\cos\theta\}-\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{u}{r^3}+\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{1}{r^2}\frac{\partial}{\partial r}\right]b_{k,u}\psi_{k,u} \\ -\sum\frac{1}{2}\frac{\mathrm{N}e^2}{mc^2}\frac{k-u}{r^3}a_{k,u}\psi_{k,u+1}.$$

We now multiply each equation by $\psi^*{}_{k',u'}$, the complex conjugate of $\psi_{k',u'}$, k' and u' having given values, and integrate through all space. The ϕ integration destroys all terms except those for which $u = u'$. The θ integration requires that k, k' should differ by unity for the terms containing $\cos\theta$, and that $k = k'$ for the remaining terms. Thus in each of the equations four terms survive, viz., three successive terms out of the first sum, the first and third containing F, and one term out of the second sum. The θ integrals are well known, and the r integrals required are

---

† See footnote, p. 315.

$$\mathrm{I}_s = \int [f_n^k(\alpha r)]^2 r^{2-s}\,dr, \qquad s = 0, 1, 2, 3,$$

$$\mathrm{J} = \int f_n^k(\alpha r) f_n^{k+1}(\alpha r)\, r^3\,dr,$$

$$\mathrm{K} = \int f_n^k(\alpha r) \frac{d}{dr} f_n^k(\alpha r)\,dr,$$

the limits 0 to ∞ being understood, and α denoting $\frac{4\pi^2 \mathrm{N} m e^2}{h^2(n+1)}$. These integrals may readily be evaluated from the definition of f_n^k by the usual method of repeated integration by parts. Thus we find

$$\mathrm{I}_0 = (n+1)\,\mathrm{I}_1 = (n+1)(k+\tfrac{1}{2})\,\mathrm{I}_2 = k(k+\tfrac{1}{2})(k+1)\,\mathrm{I}_3$$
$$= 2(n+k+2)^{-1}\,\mathrm{J} = 2^{-2n-2}(n+1)(n+k+1)!\,(n-k)!,$$
$$\mathrm{K} = -\tfrac{1}{2}[f_n^k(0)]^2.$$

The term in d/dr which gives rise to the integral K did not occur in Darwin's original equations. We now see that it is without effect when $k \neq 0$, for then $f_n^k(0) = 0$; but when $k = 0$, we have $f_n^0(0) = 2^{-2n}(n+1)!$

The equations which result are

$$-\frac{n-k+1}{2k-1}\mathrm{G}a_{k-1,u} + (\overline{\mathrm{W}} - \mathrm{W}_1 - \beta u - \mu\delta)\,a_{k,u}$$
$$-\frac{(k+u+1)(k-u+1)(n+k+2)}{2k+3}\mathrm{G}a_{k+1,u} - (k+u+1)\,\beta b_{k,u+1} = 0, \tag{3.1}$$

$$-\frac{n-k+1}{2k-1}\mathrm{G}b_{k-1,u} + (\overline{\mathrm{W}} - \mathrm{W}_1 + \beta u - \mu\delta)\,b_{k,u}$$
$$-\frac{(k+u+1)(k-u+1)(n+k+2)}{2k+3}\mathrm{G}b_{k+1,u} - (k-u+1)\,\beta a_{k,u-1} = 0, \tag{3.2}$$

where, the factor δ has the value unity when $k = 0$, otherwise it vanishes, and

$$\mathrm{W}_1 = \mu\left(\tfrac{3}{4}\frac{1}{n+1} - \frac{1}{k+\frac{1}{2}}\right),$$

$$\beta = \frac{\mu}{k(2k+1)(k+1)},$$

$$\mathrm{G} = \frac{3(n+1)h^2\mathrm{F}}{8\pi^2\mathrm{N}em},$$

$$\mu = \frac{8\pi^4\mathrm{N}^4e^8m}{c^2h^4(n+1)^3}.$$

We shall find it convenient to write Ω for $\overline{\mathrm{W}} - \tfrac{3}{4}\frac{\mu}{n+1}$.

When $k = 0$, we have $u = 0$, and the terms βu do not appear at all. These equations of course reduce to Darwin's (except for the term in δ) when $F = 0$.

In equations (3.1) and (3.2) n has a given value, and k, u assume all pairs of values subject to $0 \leqslant k \leqslant n$, $-k \leqslant u \leqslant k$; any coefficient a or b whose second suffix is numerically greater than the first vanishes, since the corresponding characteristic function does not exist. Thus we have a set of $2(n+1)^2$ linear homogeneous equations in as many variables a and b. The determinant of the set must vanish for consistency, and this gives an equation to determine the perturbations $\overline{W}$ of the energy levels due to the electric field. To each value of $\overline{W}$ corresponds a set of ratios of the coefficients a and b, from which the wave-functions f and g may be written down.

§ 4. As an example of the application of the equations, consider the case $n = 0$, which illustrates the terms in δ, but is otherwise trivial. The only admissible value of k and u is 0, so that the equations reduce to

$$(\overline{W} - W_1 - \mu)\, a_{00} = 0, \qquad (\overline{W} - W_1 - \mu)\, b_{00} = 0\,;$$

writing Ω for $\overline{W} - \frac{3}{4}\mu$, we have $\Omega + \mu = 0$, giving $\Omega = -\mu$ for the fine-structure, whether the electric field vanishes or not.

As a further illustration, and also to make the general argument, which will be outlined later, easier to follow, we shall now work through the case $n = 1$, $N = 1$ (first quantum state of hydrogen) in detail. Here

$$\Omega = \overline{W} - \tfrac{3}{8}\mu, \qquad \mu = \frac{\pi^4 m e^8}{c^2 h^4}, \qquad G = \tfrac{3}{4}\frac{h^2 F}{\pi^2 m e}.$$

We have first of all to write down the set of equations derived from (3.1) and (3.2) by putting $k = 0$, $u = 0$; $k = 1$, $u = -1, 0, 1$. They are

$$\begin{array}{cccl}
(\Omega + \frac{1}{2}\mu)\, a_{11} & & & = 0 \\
(\Omega + \mu)\, a_{00} & -\, G a_{10} & & = 0 \\
-\, G a_{00} & +\, (\Omega + \frac{2}{3}\mu)\, a_{10} & -\, \frac{1}{3}\mu b_{11} & = 0 \\
 & -\, \frac{1}{6}\mu a_{10} & +\, (\Omega + \frac{5}{6}\mu)\, b_{11} & = 0 \\
(\Omega + \mu)\, b_{00} & -\, G b_{10} & & = 0 \\
-\, G b_{00} & +\, (\Omega + \frac{2}{3}\mu)\, b_{10} & -\, \frac{1}{3}\mu a_{1-1} & = 0 \\
 & -\, \frac{1}{6}\mu b_{10} & +\, (\Omega + \frac{5}{6}\mu)\, a_{1-1} & = 0 \\
(\Omega + \frac{1}{2}\mu)\, b_{1-1} & & & = 0.
\end{array}$$

The equations have been written in this order so as to show that they fall into groups or " chains " of 1, 3, 3, 1 members, the coefficients of each chain being peculiar to it. The case $n = 1$ here illustrated is exceptional in that there are never more than three terms in an equation, instead of the full number, four. This is because the greatest number of successive a's or b's occurring in any equation is two. The determinants of the chains are identical in pairs, being

$$\Omega + \tfrac{1}{2}\mu, \qquad \begin{vmatrix} \Omega + \mu & -\mathrm{G} & 0 \\ -\mathrm{G} & \Omega + \tfrac{2}{3}\mu & -\tfrac{1}{3}\mu \\ 0 & -\tfrac{1}{3}\mu & \Omega + \tfrac{5}{6}\mu \end{vmatrix}$$

Equating these to zero, we have, to determine Ω, and hence the perturbations, the two equations

$$\Omega + \tfrac{1}{2}\mu = 0, \qquad (\Omega + \mu)^2 (\Omega + \tfrac{1}{2}\mu) = \mathrm{G}^2 (\Omega + \tfrac{5}{6}\mu).$$

When $\mathrm{G} = 0$, these give the two fine-structure levels $\Omega = -\tfrac{1}{2}\mu,\ -\mu$; and when $\mathrm{G} \neq 0$, the first still gives $\Omega = -\tfrac{1}{2}\mu$, while the second gives, in first approximation, when $\mathrm{G} \ll \mu$ ("weak" fields),

$$\Omega = -\tfrac{1}{2}\mu + \tfrac{4}{3}\frac{\mathrm{G}^2}{\mu}, \qquad -\mu \pm \frac{\mathrm{G}}{\sqrt{3}}.$$

To each root corresponds a set of ratios of the eight coefficients a and b. Thus when $\Omega = -\tfrac{1}{2}\mu$, the first chain gives $a_{11} =$ const., and when $\Omega = -\tfrac{1}{2}\mu + \tfrac{4}{3}\frac{\mathrm{G}^2}{\mu}$, the second chain gives $a_{00} : a_{10} : b_{11} :: 0 : 2 : 1$, the other coefficients in each case all vanishing.*

In "strong" fields ($\mathrm{G} \gg \mu$) the determinants are, if we neglect μ altogether,

$$\begin{vmatrix} \Omega & -\mathrm{G} & 0 \\ -\mathrm{G} & \Omega & 0 \\ 0 & 0 & \Omega \end{vmatrix}$$

so that the null approximations to the levels are $\Omega = 0,\ \pm \mathrm{G}$. We remark in passing that the third order determinant splits up into the product of two principal minors, a property which will be found to hold in the general case. If we retain the first power of μ, the first approximations to the levels are $\Omega = -\tfrac{5}{6}\mu,\ \pm \mathrm{G} - \tfrac{5}{6}\mu$. The ratios of the a's and b's are determined exactly as before. The results are given in Table I, which gives the levels and $a : b$ ratios for weak and strong fields for the first quantum state of hydrogen. The levels are given in first approximation, while the $a : b$ ratios are, strictly speaking, the null approximation corresponding to zero perturbation in each case.

---

* The coefficients of the other chain of a pair having the same determinant may be supposed to vanish because the residual degeneracy which the identity of the determinants implies may be removed, *e.g.*, by a small magnetic perturbation.

Table I.—$N = 1$, $n = 1$.

State.			Weak fields.			Strong fields.		
m	j	k	Ω	a's	b's	Ω	a's	b's
$\frac{3}{2}$	$\frac{3}{2}$	1	$-\frac{1}{2}\mu$	1		$-\frac{1}{2}\mu$	1	
	$\frac{3}{2}$	1	$-\frac{1}{2}\mu + \frac{4}{3}\frac{G^2}{\mu}$	0 2	1	$G - \frac{5}{6}\mu$	1 1	0
$\frac{1}{2}$	$\frac{1}{2}$	1	$-\mu + \frac{G}{\sqrt{3}}$	$\sqrt{3}$ 1	-1	$-\frac{5}{6}\mu$	0 0	1
	$\frac{1}{2}$	0	$-\mu - \frac{G}{\sqrt{3}}$	$-\sqrt{3}$ 1	-1	$-G - \frac{5}{6}\mu$	1 -1	0
	$\frac{3}{2}$	1	$-\frac{1}{2}\mu + \frac{4}{3}\frac{G^2}{\mu}$	1	0 2	$G - \frac{5}{6}\mu$	0	1 1
$-\frac{1}{2}$	$\frac{1}{2}$	1	$-\mu + \frac{G}{\sqrt{3}}$	-1	$\sqrt{3}$ 1	$-\frac{5}{6}\mu$	1	0 0
	$\frac{1}{2}$	0	$-\mu - \frac{G}{\sqrt{3}}$	-1	$-\sqrt{3}$ 1	$-G - \frac{5}{6}\mu$	0	1 -1
$-\frac{3}{2}$	$\frac{3}{2}$	1	$-\frac{1}{2}\mu$		1	$-\frac{1}{2}\mu$		1

$$\Omega = \overline{W} - \tfrac{5}{8}\mu, \qquad \mu = \frac{\pi^4 m e^8}{c^2 h^4}, \qquad G = \tfrac{3}{4}\,\frac{h^2 F}{\pi^2 m e}.$$

"Weak" and "strong" levels written on the same horizontal line have been associated together by the principle that as G increases the roots of a chain remain real, and therefore cannot cross, so that their serial order, from highest to lowest, does not alter. Thus, for instance, the roots

$$-\tfrac{1}{2}\mu + \tfrac{4}{3}\frac{G^2}{\mu}, \quad -\mu + \frac{G}{\sqrt{3}}, \quad -\mu - \frac{G}{\sqrt{3}},$$

which the second or third chain gives in weak fields, become

$$G - \tfrac{5}{6}\mu, \quad -\tfrac{5}{6}\mu, \quad -G - \tfrac{5}{6}\mu$$

respectively in strong fields. The calculation of levels and $a : b$ ratios has also been carried out for the second quantum state of hydrogen, the results being shown in Table II. To save space, only the first half of the table is given; the other half may be completed without difficulty on the model of Table I, which is given in full. The meaning of the quantum numbers j, k, m, will be

Table II.—N = 1, $n = 2$.

(Only the first half of the table is given.)

State.			Weak fields.						Strong fields.					
m	j	k	Ω	a's.			b's.		Ω	a's.			b's.	
$\frac{5}{2}$	$\frac{5}{2}$	2	$-\frac{1}{3}\mu$	1					$-\frac{1}{3}\mu$	1				
	$\frac{5}{2}$	2	$-\frac{1}{3}\mu + \frac{24}{5}\frac{G^2}{\mu}$	0	4		1		$G - \frac{13}{30}\mu$	3	1		0	
$\frac{3}{2}$	$\frac{3}{2}$	2	$-\frac{1}{2}\mu + \frac{G}{\sqrt{5}}$	$3\sqrt{5}$	1		-1		$-\frac{14}{30}\mu$	0	0		1	
	$\frac{3}{2}$	1	$-\frac{1}{2}\mu - \frac{G}{\sqrt{5}}$	$-3\sqrt{5}$	1		-1		$-G - \frac{13}{30}\mu$	3	-1		0	
	$\frac{5}{2}$	2	$-\frac{1}{3}\mu + \frac{36}{5}\frac{G^2}{\mu}$	0	0	3	0	2	$2G - \frac{22}{30}\mu$	4	6	1	0	0
	$\frac{3}{2}$	2	$-\frac{1}{2}\mu + \frac{G}{3\sqrt{5}}$	0	$2\sqrt{5}$	1	$\sqrt{5}$	-1	$G - \frac{19}{30}\mu$	0	0	0	3	1
$\frac{1}{2}$	$\frac{3}{2}$	1	$-\frac{1}{2}\mu - \frac{G}{3\sqrt{5}}$	0	$-2\sqrt{5}$	1	$-\sqrt{5}$	-1	$-\frac{18}{30}\mu$	2	0	-1	0	0
	$\frac{1}{2}$	1	$-\mu + \frac{2\sqrt{2G}}{3}$	$\sqrt{2}$	1	0	-1	0	$-G - \frac{19}{30}\mu$	0	0	0	-3	1
	$\frac{1}{2}$	0	$-\mu - \frac{2\sqrt{2G}}{3}$	$-\sqrt{2}$	1	0	-1	0	$-2G \quad \frac{22}{30}\mu$	4	-6	1	0	0

$$\Omega = \overline{W} - \tfrac{1}{4}\mu, \qquad \mu = \tfrac{8}{27}\frac{\pi^4 me^8}{c^2h^4}, \qquad G = \tfrac{9}{8}\frac{h^2F}{\pi^2 me}.$$

explained later. We shall return to these tables in working out the intensities in the line H_a.

The two diagrams in fig. 1 show the levels in weak fields (above) and strong fields (below), together with the connections between them, for the two cases $n = 1$, $n = 2$. No attempt has been made to draw the diagrams to scale; they are intended merely to help to visualise the information in the tables. The uppermost level in weak fields (on the extreme right) is actually double in $n = 1$ and triple in $n = 2$. Since the separation is proportional to the square of the field, the individual components have not been shown, but their existence has been indicated by drawing the transition lines branching out from the levels in question curved instead of straight.

§ 5. We shall now consider briefly the general properties of the system of equations derived from (3.1) and (3.2) by allowing k and u to assume their admissible values. Each equation of the system involves only two suffixes u, differing by unity, the suffix which is algebraically less being common to all the a's and the other to the b's. We shall denote the mean value of the two suffixes u occurring in any equation by m, and group together in a chain all the equations having the same value of m, which accordingly serves to identify a chain.

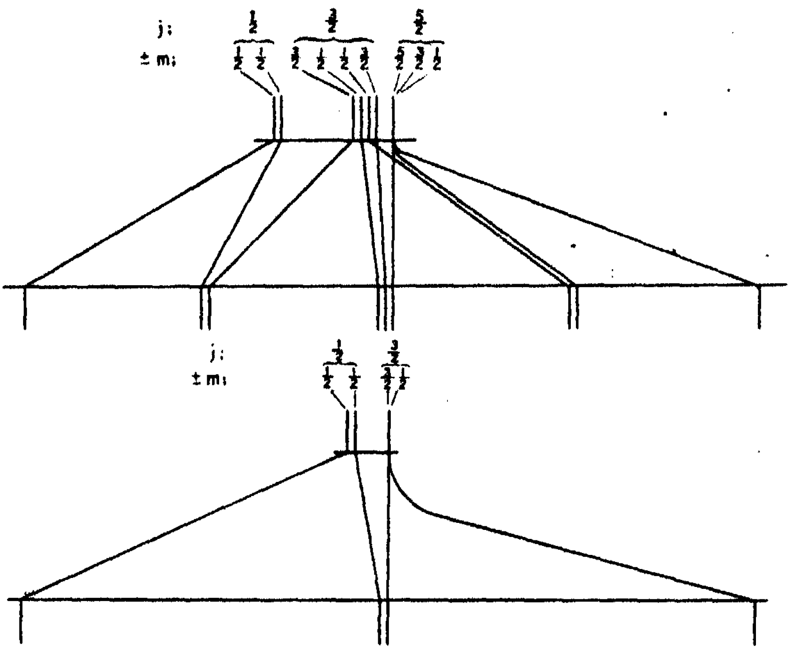

j: $\frac{1}{2}$ $\frac{3}{2}$ $\frac{5}{2}$
± m: $\frac{1}{2}$ $\frac{1}{2}$ $\frac{3}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{3}{2}$ $\frac{5}{2}$ $\frac{3}{2}$ $\frac{1}{2}$
j: $\frac{1}{2}$ $\frac{3}{2}$
± m: $\frac{1}{2}$ $\frac{1}{2}$ $\frac{3}{2}$ $\frac{1}{2}$

FIG. 1.—Diagrams of Transitions between Levels in Weak and Strong Fields for First and Second Quantum States.

From the equations it is evident that the a's and b's in the chain m are

$$a_{|m-\frac{1}{2}|,\, m-\frac{1}{2}}, \quad a_{|m-\frac{1}{2}|+1,\, m-\frac{1}{2}}, \quad a_{|m-\frac{1}{2}|+2,\, m-\frac{1}{2}} \;\cdots\cdots\; a_{n,\, m-\frac{1}{2}}$$
$$b_{|m+\frac{1}{2}|,\, m+\frac{1}{2}}, \quad b_{|m+\frac{1}{2}|+1,\, m+\frac{1}{2}}, \quad b_{|m+\frac{1}{2}|+2,\, m+\frac{1}{2}} \;\cdots\cdots\; b_{n,\, m+\frac{1}{2}}.$$

Hence the highest value m can assume is $n + \frac{1}{2}$, and the lowest $-(n + \frac{1}{2})$, the corresponding chains having only one member each. The chains $m = \pm (n - \frac{1}{2})$ each have three members, $m = \pm (n - \frac{3}{2})$ five each, and so on, the number of members in a chain increasing by two at a time until the central chains, $m = \pm \frac{1}{2}$, having $2n + 1$ members each, are reached.

Consider now the relation between the chain m and the chain $-m$. The former is generated by writing $u = m - \frac{1}{2}$ in (3.1) and $u = m + \frac{1}{2}$ in (3.2),

k in both cases being supposed to run through its admissible values. Comparing equations (3.1) and (3.2) using these values of u, we see that the coefficients of any of the quantities $a_{p,q}$, $b_{r,s}$ in the chain m are respectively equal to the coefficients of $b_{p,-q}$, $a_{r,-s}$ in the chain $-m$. The determinants of the two chains $\pm m$ are therefore identical, and yield the same set of roots. Thus the levels $\pm m$ fall together at all strengths of field, so that there are $(n+1)^2$ distinct levels instead of the full number $2(n+1)^2$, which would arise if the system were completely non-degenerate, as, for instance, in the Zeeman effect. We may express this by saying that the electric field, unlike the magnetic field, does not discriminate between right-handed and left-handed orbits.

In the absence of an electric field $G = 0$, and the chain m consists of the equations

$$\begin{gathered}
\{\overline{W} - W_1 - \beta(m - \tfrac{1}{2})\}\, a_{k,m-\frac{1}{2}} - (k + m + \tfrac{1}{2})\,\beta\, b_{k,m+\frac{1}{2}} = 0, \\
-(k - m + \tfrac{1}{2})\,\beta\, a_{k,m-\frac{1}{2}} + [\overline{W} - W, + \beta(m + \tfrac{1}{2})]\, b_{k,m-\frac{1}{2}} = 0,
\end{gathered}$$

where k has the values $m - \frac{1}{2}$, $m + \frac{1}{2}$, $m + \frac{3}{2}$ …… n in succession. When $k = m - \frac{1}{2}$, the second of these is satisfied identically, while the first gives

$$\{\overline{W} - W_1 - \beta(m - \tfrac{1}{2})\}\, a_{m-\frac{1}{2},m-\frac{1}{2}} = 0,$$

whence

$$\Omega = -\frac{\mu}{m + \frac{1}{2}}.$$

For each of the remaining values of k there is a pair of equations whose roots are

$$\Omega = -\frac{\mu}{k}, \quad \Omega = -\frac{\mu}{k+1},$$

where

$$k = m + \tfrac{1}{2}, \quad m + \tfrac{3}{2}, \; \ldots\ldots \; n.$$

The determinant of the chain m may accordingly be written, in the absence of an electric field,

$$\left(\Omega + \frac{\mu}{m - \frac{1}{2}}\right)^2 \left(\Omega + \frac{\mu}{m + \frac{1}{2}}\right)^2 \ldots\ldots \left(\Omega + \frac{\mu}{n}\right)^2 \left(\Omega + \frac{\mu}{n+1}\right) = 0,$$

which shows that when $G = 0$ the chain m contains the highest fine-structure level as a single root, and the next $n - m - \frac{1}{2}$ lower each as a double root.

Suppose now that there is an electric field, so that G no longer vanishes. We shall show that the chain determinants are then even functions of G. Consider the equations (3.1) and (3.2). To obtain a chain from them, we must write $u+1$ for u in (3.2) and allow k to run through its admissible values. The coefficients in the equations of the chain will be

$$\begin{array}{cccc} a_{k-1,\,u} & a_{k,\,u} & a_{k+1,\,u} & b_{k,\,u+1} \\ b_{k-1,\,u+1} & b_{k,\,u+1} & b_{k+1,\,u+1} & a_{k,\,u} \end{array}$$

where u is fixed and k varies. Now change the sign of alternate coefficients, viz.,

$$\begin{array}{ccc} \ldots\ldots a_{k-1,\,u} & a_{k+1,\,u} & a_{k+3,\,u}\ldots\ldots \\ \ldots\ldots b_{k-1,\,u+1} & b_{k+1,\,u+1} & b_{k+3,\,u+1}\ldots\ldots \end{array}$$

and at the same time write $-\mathrm{G}$ for G. Then (3.1) and (3.2) return to their original form. Hence the two equations for which $k=k'$, say, are unaffected, and the same applies to those for which $k=k'\pm 2, k'\pm 4, \ldots$. The alternate equations for $k=k'\pm 1$, $k'\pm 3$, ... have their sign reversed throughout. The number of equations in the whole chain which are thus affected is, of course, even.

Hence the double operation of changing the sign of alternate a's and b's and writing $-\mathrm{G}$ for G is without effect on the value of the determinant of the chain, being equivalent merely to changing the sign of an even number of rows. But the first operation alone is equivalent to changing the sign of an even number of columns, and is therefore without effect on the determinant; consequently the same must be true of the second alone. In other words, the chain determinant is an even function of G. This property of the determinants is evidently associated with the fact that reversal of the field must give rise to the same set of energy levels, the groups of levels $+m$ and $-m$ being merely interchanged. Since these are roots of identical determinants, the latter must be independent of reversal of field. Note that this argument would not apply to the Zeeman effect, because there the levels $+m$ and $-m$ are distinct.

We have seen that the expansion of the determinant involves even powers of G only; we shall show later* that there is a term in G^2, so that when the field is so weak that higher powers of G may be neglected, we may write, for the determinant of the chain m,

$$\left(\Omega+\frac{\mu}{m-\frac{1}{2}}\right)^2\left(\Omega+\frac{\mu}{m+\frac{1}{2}}\right)^2\ldots\left(\Omega+\frac{\mu}{n}\right)^2\left(\Omega+\frac{\mu}{n+1}\right)=\mathrm{G}^2 f(\Omega,\mu),$$

where the function f does not contain G; from which it immediately follows that the perturbation of the single root $\Omega=-\dfrac{\mu}{n+1}$ is proportional to G^2

---

* *Vide* Note at the end of this paper.

in first approximation, and the separation of the roots which coincide in the absence of a field is symmetrical and proportional to $\pm$ G. Each value of m will give rise to a different set of perturbations of the levels, which may be calculated from the appropriate determinant. The difference in the behaviour of the highest level does not arise in the Zeeman effect, because there levels which coincide in the absence of a field occur as roots of different chains.

We now go on to consider the other extreme case, in which the electric field is so strong that its effect is much more important than the fine-structure. If F is so large that the terms in $1/c^2$ in (3.1) and (3.2) may be neglected altogether in comparison with those in G, the two central chains must, of course, give as roots of their determinant all the levels $\overline{\mathrm{W}} = \pm s\mathrm{G}$ (s integral, and $0 \leqslant s \leqslant n$), originally found to this degree of approximation by Schwarzschild and Epstein. The proof that our chain determinants actually give these levels is outlined in a note appended to this paper. We may readily calculate, in any particular case, the first approximations to the levels when the fine-structure terms in $1/c^2$ are taken into account. The displacements of the two levels $\pm s\mathrm{G}$ calculated from a given chain will be equal and in the same sense, since the determinants are even functions of G; but different chains give different displacements. Hence the central Stark level, $\overline{\mathrm{W}} \doteqdot 0$, which is a root of all the chains, possesses $(n + 1)$ fine-structure levels, the next two outside, $\overline{\mathrm{W}} \doteqdot \pm \mathrm{G}$, which are roots of all but the two end chains, possess n each, and so on, the two extreme levels remaining single.

As regards the coefficients a and b, it is apparent from the structure of the determinants that either the a's or the b's associated with a given level all vanish, so that here the characteristic functions are either f's or g's, whereas in weak fields there is in general both an f and a g associated with each level.

§ 6. We have now to consider a suitable choice of three quantum numbers to specify the state of the atom for a given value of n. As usual, we may number the unperturbed fine-structure levels $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2} \ldots n + \frac{1}{2}$ from the lowest to the highest. In weak fields all the levels into which any of the fine-structure levels breaks up have the same value of j. The quantum number m and its relation to u have been explained above. The remaining quantum number is conveniently taken to be k. The roots derived from a given chain, that is, the levels having a given value of m, will be associated with the several suffixes k occurring in that chain. In the absence of an electric field it is an easy matter to see how the values of k are distributed among the roots; we are led in this way to the Sommerfeld classification, according to which $j = \frac{1}{2}$ has $k = 0, 1$, $j = \frac{3}{2}$ has $k = 1, 2$, and so on, $j = n$ has $k = n - 1, n$, and finally

$j = n + \frac{1}{2}$ has $k = n$ only. In the presence of a weak electric field k ceases to be a quantum number, but we can still conventionally assign its two values to the two solutions having the same j and m. We shall suppose that the greater of the two possible values of k is assigned to the higher level of the pair. The quantum numbers in Tables I and II have been written down according to this scheme. If we arrange the roots of any chain in weak fields in order of magnitude, this order is preserved as the roots alter under the influence of an increasing field. The connection between the levels in weak fields and those in strong fields may be expressed by the statement that the level $(n\,j\,m\,k)$ becomes $(k + j - n - |m|)$ G in strong fields if we ignore the fine-structure. The combination $k + j - n - |m|$ (here denoted by s) is thus the difference between the two " parabolic " quantum numbers of the older analysis.

§ 7. It would be very troublesome to obtain general algebraic formulæ for the intensity of any line, so we shall content ourselves with discussing the intensities of the components of H$_\alpha$ (6562 Å), which is given by the transition from $n = 2$ to $n = 1$; the same method is, of course, applicable to any other hydrogen-like line. In the present instance Tables I and II enable us to write down the characteristic functions for any of the levels of either state, in weak or strong fields. The coefficients a and b in the chain m have been written down above; the numbers in the columns of Tables I and II headed " a's b's " give the ratio of these coefficients, in that order, for different values of j and k. Thus, for instance, when $(n\,j\,m\,k) = (1\,\frac{3}{2}\,\frac{1}{2}\,1)$, the ratios of $a_{00} : a_{10} : b_{11}$ are $0 : 2 : 1$ in weak fields and $1 : 1 : 0$ in strong fields. The characteristic functions

$$f = a_{00}\,\mathrm{P}_0{}^0 f_1{}^0 + a_{1-1}\,\mathrm{P}_1{}^{-1} f_1{}^1 e^{-i\phi} + a_{10}\,\mathrm{P}_1{}^0 f_1{}^1 + a_{11}\,\mathrm{P}_1{}^1 f_1{}^1 e^{i\phi}$$

$$g = b_{00}\,\mathrm{P}_0{}^0 f_1{}^0 + b_{1-1}\,\mathrm{P}_1{}^{-1} f_1{}^1 e^{-i\phi} + b_{10}\,\mathrm{P}_1{}^0 f_1{}^1 + b_{11}\,\mathrm{P}_1{}^1 f_1{}^1 e^{i\phi}$$

are therefore $2\,\alpha r e^{-\alpha r}\cos\theta$ and $\alpha r e^{-\alpha r}\sin\theta$ respectively in weak fields, and $(\alpha r - 1)\,e^{-\alpha r} + \alpha r e^{-\alpha r}\cos\theta$ and 0 respectively in strong fields, where $\alpha = 2\pi^2 m e^2/h^2$.

The intensity of the line corresponding to a given transition between the levels of two different quantum states n_p, n_q, is given[†] by

$$\frac{\left|\int\left\{\begin{matrix} x - iy \\ 2z \\ x + iy \end{matrix}\right\}(f_p f_q{}^* + g_p g_q{}^*)\,dx\,dy\,dz\right|^2}{\int(|f_p|^2 + |g_p|^2)\,dx\,dy\,dz \,.\, \int(|f_q|^2 + |g_q|^2)\,dx\,dy\,dz},$$

the asterisk denoting the complex conjugate. The type of polarisation is determined by the factor in the numerator. Since each characteristic function is associated with a definite value of m, the f's will contain a factor $e^{i(m-\frac{1}{2})\phi}$, and the g's $e^{i(m+\frac{1}{2})\phi}$. Since f_p and f_q^*, also g_p and g_q^* are multiplied together, the only transitions for which the intensities do not vanish are those in which m changes by -1, 0, $+1$, the polarisations of the corresponding lines being respectively $\perp l$, $\parallel$, $\perp r$. Since the levels $\pm m$ coincide, a line for which $m_p \rightarrow m_q$ will coincide with one for which $-m_p \rightarrow -m_q$. The intensities of the two lines will be equal, and their polarisation will be either $\parallel$ for both, so that they disappear on longitudinal observation, or $\perp l$ for one and $\perp r$ for the other, so that observed longitudinally the line is unpolarised. The case $m = \pm \frac{1}{2} \rightarrow \pm \frac{1}{2}$ is particularly interesting, since here the possibility of partial polarisation arises, because $\frac{1}{2} \rightarrow \frac{1}{2}$ gives $\parallel$ and $\frac{1}{2} \rightarrow -\frac{1}{2}$ $\perp$ polarisation. It may, of course, happen (as, for example, in H_a in strong fields) that one or other of these components vanishes, so that the polarisation is complete.

[†] Darwin *l.c.*, p. 241.

§ 8. The intensities of all the components of H_a in weak and strong fields have been calculated from the above formula, the results being given in Tables III and IV. Before discussing these tables we shall describe briefly the general results.

In the absence of an electric field the state $n = 1$ has two fine-structure levels and $n = 2$ has three. Sommerfeld and Unsöld have shown that the six fine-structure components arising from these have intensities in the ratio 0 : 21760 : 3500 : 27648 : 3072 : 600, the component whose intensity vanishes corresponding to the transition $j = \frac{5}{2} \rightarrow \frac{1}{2}$. We have seen that the first effect of a weak electric field F is to split up each of the fine-structure levels, the separation in the highest levels of both states being proportional to F^2, and in the others proportional to F itself. The upper level of each state may therefore be regarded, relatively to the others, as remaining single in weak fields. The present calculation shows that the group of lines into which any fine-structure component is split by a weak field is unpolarised on the whole, and has a total intensity agreeing exactly with the ratios given above. In particular the transition $j = \frac{5}{2} \rightarrow \frac{1}{2}$ gives no lines; this will be true so long as the null approximation to the a's and b's is valid, but in fields which are no longer quite weak it is necessary to use a higher approximation, and it is evident that the prohibited line will then begin to appear with an intensity proportional to the square or other even power of F.

It will be remembered that Kramers' theory leads to a Stark effect of the fine-structure components in which the separation is proportional to F^2 in weak fields; according to the present theory, however, the separation in some of

Table III.

$j'_{\cdot}\rightarrow j$.	σ.	τ.	I $\parallel$	I $\perp$	Σ I	No. in strong fields.
$\frac{1}{2}$ $\frac{3}{2}$	−32	−1·06	50	50		1 2 3
		1·06	50	50	200	6 7 11
$\frac{3}{2}$ $\frac{3}{2}$	−8	−0·50	230	77		8 12
		−0·17	26	179		16 17 22
		0·17	26	179		28 29 34
		0·50	230	77	1024	18 23
$\frac{5}{2}$ $\frac{3}{2}$	0	0·00	4608	4608	9216	24 30 35 39 40 44
$\frac{1}{2}$ $\frac{1}{2}$	49	−1·49	64	277		4 5
		−0·63	227	64		9 10
		0·63	227	64		13 14
		1·49	64	277	3500/3	19 20
$\frac{3}{2}$ $\frac{1}{2}$	73	−0·94	0	60		15
		−0·60	80	433		25 26
		−0·27	0	1300		21
		−0·07	1733	20		36 37
		0·07	1733	20		31 32
		0·27	0	1300		27
		0·60	80	433		41 42
		0·94	0	60	21760/3	33
$\frac{5}{2}$ $\frac{1}{2}$	81	−0·43	0	0		38 45 46
		0·43	0	0	0	43 47 48

the fine-structure components is proportional to the first power of the field. This affords a possibility of experimental test, but it would probably be very difficult to measure an incipient Stark effect in fine-structure components which it is scarcely possible to resolve in the absence of a field. We should, however, as the figures in the Table III show, expect the edges of the fine-structure lines to be differently polarised from the centres in presence of a weak field; thus the outer edges of $j=\frac{1}{2}\rightarrow\frac{1}{2}$ and $\frac{3}{2}\rightarrow\frac{1}{2}$ are predominantly $\perp$, and those of $j=\frac{3}{2}\rightarrow\frac{3}{2}$ predominantly $\parallel$. This effect is perhaps more accessible to observation.

Table IV.

No.	2m′ 2j′ k′	2m 2j 2k	σ′	τ′	P	I	ΣI
1	1 1 0	−1 3 1	−8	87	⊥	0	
2	1 1 0	1 3 1		87	‖	1	1
3	1 1 0	3 3 1	−6	− 48	⊥	9	
4	1 1 0	−1 1 1		87	⊥	9	
5	1 1 0	1 1 1		87	‖	0	18
6	1 1 1	−1 3 1	−5	99	⊥	8	
7	1 1 1	1 3 1		99	‖	0	
8	3 3 1	1 3 1		123	⊥	8	16
9	1 1 0	−1 1 0	−4	87	⊥	0	
10	1 1 0	1 1 0		87	‖	1681	1681
11	1 1 1	3 3 1	−3	− 36	⊥	0	
12	3 3 1	3 3 1		− 12	‖	1152	
13	1 1 1	−1 1 1		99	⊥	0	
14	1 1 1	1 1 1		99	‖	1152	
15	3 3 1	1 1 1		173	⊥	0	2304
16	1 3 1	−1 3 1	−2	103	⊥	0	
17	1 3 1	1 3 1		103	‖	729	
18	3 3 2	1 3 1		119	⊥	0	729
19	1 1 1	−1 1 0	−1	99	⊥	968	
20	1 1 1	1 1 0		99	‖	0	
21	3 3 1	1 1 0		123	⊥	968	1936
22	1 3 1	3 3 1	0	− 32	⊥	441	
23	3 3 2	3 3 1		− 16	‖	0	
24	5 5 2	3 3 1		0	⊥	2304	
25	1 3 1	−1 1 1		103	⊥	441	
26	1 3 1	1 1 1		103	‖	0	
27	3 3 2	1 1 1		119	⊥	2304	5490
28	1 3 2	−1 3 1	1	99	⊥	968	
29	1 3 2	1 3 1		99	‖	0	
30	3 5 2	1 3 1		123	⊥	968	1936
31	1 3 1	−1 1 0	2	103	⊥	0	
32	1 3 1	1 1 0		103	‖	729	
33	3 3 2	1 1 0		119	⊥	0	729
34	1 3 2	3 3 1	3	− 36	⊥	0	
35	3 5 2	3 3 1		− 12	‖	1152	
36	1 3 2	−1 1 1		99	⊥	0	
37	1 3 2	1 1 1		99	‖	1152	
38	3 5 2	1 1 1		173	⊥	0	2304
39	1 5 2	−1 3 1	4	87	⊥	0	
40	1 5 2	1 3 1		87	‖	1681	1681
41	1 3 2	−1 1 0	5	99	⊥	8	
42	1 3 2	1 1 0		99	‖	0	
43	3 5 2	1 1 0		123	⊥	8	16
44	1 5 2	3 3 1	6	− 48	⊥	9	
45	1 5 2	−1 1 1		87	⊥	9	
46	1 5 2	1 1 1		87	‖	0	18
47	1 5 2	−1 1 0	8	87	⊥	0	
48	1 5 2	1 1 0		87	‖	1	1

In strong fields we have shown that each of the levels $\pm sG$ of either state has a fine-structure on the same scale as the fine-structure of the original levels in the absence of a field. This fine-structure will be reproduced in the Stark components. The approximation which neglects the fine-structure leads to components whose displacements from the position of the leading fine-structure component without field ($j = \frac{5}{2} \rightarrow \frac{3}{2}$) are proportional to $\pm$ (01234568), of which $\pm$ (0156) are $\perp$ and $\pm$ (2348) are $\|$. Their intensities, calculated by Schrödinger, are in the ratio 2745 : 1936 : 729 : 2304 : 1681 : 16 : 18 : 1. (The value 2745 for the undisplaced component (0) has been halved; the intensities of the other components are symmetrical about it.) According to the present analysis, each of these lines has several fine-structure components; but the number of these components whose intensity differs from zero is considerably reduced by the vanishing of alternate f's and g's in Tables I and II. Thus the Stark components $\pm$ (248) are single, $\pm$ (1356) double, and (0) quadruple. The two components of each of $\pm$ (1356) are of equal intensity; in general, the component $m' \rightarrow m$ has the same intensity as $m' - 1 \rightarrow m - 1$ provided both belong to the same Stark component. This is a consequence of the relation* between sub-chains of successive chains, and is easily verified from the tables. The surviving fine-structure components are no longer symmetrical about the positions $\pm$ (01234568), nor is the whole complex of lines symmetrical about (0).

The rule of combination for the quantum number m has already been pointed out. It is perfectly general, and applies in weak and strong fields. In the case of H_α we have shown that in weak fields the component for which j changes by two units ($j = \frac{5}{2} \rightarrow \frac{1}{2}$) has zero intensity. This illustrates the general combination rule that the intensity of a component for which j changes by more than one unit has an intensity proportional to the square or other even power of the field. In strong fields the rule, of course, breaks down. In weak fields k is purely conventional; no combination rule exists for it. In strong fields we have the combination rule due to the vanishing of alternate f's and g's. We may express this by the statement that in the transition $(n'j'm'k') \rightarrow (n\,j\,m\,k)$ the four differences $j' - j$, $m' - m$, $k' - k$, $|m'| - |m|$ must either be all odd or all even, or two even and two odd, otherwise the intensity vanishes. This may readily be verified from Table IV.

§ 9. The numerical results of the calculation are given, along with the positions of the lines, in Tables III and IV. It has not been thought worth while to give the full table for weak fields, since many of the separations are far too small to be observed. In Table III we have accordingly treated as single any group of components having a separation proportional to the square of the field. The first column in Table III gives the transitions of the quantum number j giving rise to the various unperturbed fine-structure components of H_α; their positions, referred to the leading component $j = \frac{5}{2} \to \frac{3}{2}$, are given in the third column, the numbers in which give the frequency difference $\Delta\nu$ in multiples of $\frac{1}{162}\frac{\pi^4 me^8}{c^2h^5}$. The fourth column gives the approximate positions of the components into which the various fine-structure lines are broken up in weak fields; in this column components whose separation is proportional to the square of the field are regarded as indistinguishable. The position of any component is to be found from the relation

---

* *Vide* Note at the end of this paper, p. 332.

$$h\Delta\nu = \frac{1}{162}\frac{\pi^4 me^8}{c^2h^4}\sigma + \frac{h^2\mathrm{F}}{\pi^2 me}\tau.$$

In the fifth column are given the approximate relative intensities of the Stark components; this column exhibits the intensities of $\parallel$ and $\perp$ radiation in the component separately, the polarisation being in general partial. Each fine-structure component is, of course, unpolarised on the whole. The sixth column gives the total intensity of the Stark components of each line of the fine-structure, and the figures are seen to agree with the results of Sommerfeld and Unsöld. The figures in this column are exact; those in the previous column, being generally irrational, have been rounded off to the nearest integer, which accounts for the small discrepancies in the sums. The numbers in the last column refer to those in the first column of Table IV, and serve to connect the components in weak fields with those in strong fields. In Table IV, which applies to strong fields, the first column is merely the ordinal number of the component when these are arranged according to frequency. The columns headed $2m'$, $2j'$, k', $2m$, $2j$, k give the quantum numbers of the two states between which a transition takes place, accented letters referring to $n = 2$ and unaccented to $n = 1$. The quantum numbers m and j have been multiplied by two merely to save unnecessary printing. The next two columns give the positions of the components relative to the leading component $(m\,j\,k) = (\frac{3}{2}\,\frac{3}{2}\,2) \to (\frac{1}{2}\,\frac{1}{2}\,1)$, which keeps its position as the field increases; the frequency difference $\Delta\nu$ is given by the formula

$$h\Delta\nu = \frac{3}{8}\frac{h^2\mathrm{F}}{\pi^2 me}\sigma' + \frac{1}{16}\frac{\pi^4 me^8}{c^2h^4}\tau',$$

the first term of which fixes the Stark component, and the second its fine-structure. The next column gives the polarisation, and the next the intensity of each line. The last column gives the total intensity of each Stark component; the figures in this column are in exact agreement with Schrödinger's results. In this table all the numbers are rational and have not been rounded off. The figures in the columns of both tables headed Σ I are comparable with each other.

Fig. 2 represents the line H_α in a weak field before the Stark components of the original fine-structure have begun to overlap and perturb each other. All the separations which appear in the figure are proportional to the first power of the field. The || components are shown above, and the $\perp$ components below the horizontal line. The lengths of the vertical lines are proportional to the intensities; but as it is not practicable to represent them all on the same scale, the lines shown dotted have been drawn one-eighth, and those which are merely interrupted by two dots one-half, of their correct length. The strongest || component of $j = \frac{3}{2} \rightarrow \frac{1}{2}$ is a close doublet; the intensity shown is its total intensity.

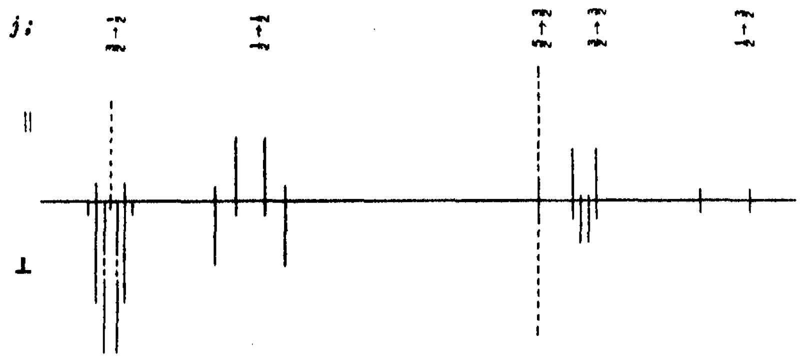


FIG. 2.—Fine-structure H_α in weak field.

Note on the Chain Determinants in Strong Fields.

In this note we outline the proof of certain properties of the chain determinants made use of in the course of the work. We shall suppose that the field is so strong that the terms in $1/c^2$ occurring in equations (3.1) and (3.2) may be neglected in comparison with those in G. Then it is evident that each chain of equations except the first and the last breaks up into two sub-chains independent of each other, one involving the coefficients a only and the other the coefficients b only. The determinant of each chain therefore breaks up into the product of two principal minors, the second of which is identical with the first of the two minors into which the determinant of the preceding chain breaks up. This follows because the terms in G in equations (3.1) and (3.2) are identical. The members of the first sub-chain of the chain $m = u + \frac{1}{2}$ $(u \geqslant 0)$ are obtained by writing $k = u,\ u+1,\ u+2 \ldots n$ in succession in equation (3.1). The determinant of this sub-chain is seen to be a continuant, which we write with its diagonals horizontal to economise space,

$$\begin{array}{ccccc} & \dfrac{-1\,(2u+1)\,(n+u+2)}{2u+3}\,\mathrm{G} & & \dfrac{-2\,(2u+2)\,(n+u+3)}{2u+5}\,\mathrm{G}\ldots & \\ \overline{\mathrm{W}} & & \overline{\mathrm{W}} & & \overline{\mathrm{W}} \\ & \dfrac{-(n-u)}{2u+1}\,\mathrm{G} & & \dfrac{-(n-u-1)}{2u+3}\,\mathrm{G}\ldots . & \end{array}$$

In a continuant it is allowable to interchange factors between conjugate elements (here written vertically below each other) provided that their product remains unchanged. In this way we obtain

$$\begin{array}{ccccc} & \dfrac{-(2u+1)\,(n-u)}{2u+1}\,\mathrm{G} & & \dfrac{-(2u+2)\,(n-u+1)}{2u+3}\,\mathrm{G}\ldots & \\ \overline{\mathrm{W}} & & \overline{\mathrm{W}} & & \overline{\mathrm{W}} \\ & \dfrac{-1\,(n+u+2)}{2u+3}\,\mathrm{G} & & \dfrac{-2\,(n+u+3)}{2u+5}\,\mathrm{G}\ldots, & \end{array}$$

which we shall denote by $\mathrm{K}_{n,u}$. To find its factors,* add to the elements of each column the sum of the corresponding elements in every alternate succeeding column. Since the rth row consists of the three adjacent elements

$$\frac{-(r-1)\,(n+u+r)}{2u+2r-1}\,\mathrm{G},\quad \overline{\mathrm{W}},\quad \frac{-(2u+r)\,(n-u-r+1)}{2u+2r-1},$$

the result of this operation is to give all the elements lying below the principal diagonal the values $\overline{\mathrm{W}}$ and $-(n-u)\,\mathrm{G}$ alternately, while those above the diagonal are unaffected. If we now subtract from each row the row next but one above it, the determinant at once breaks up into the product of the second order minor

$$\begin{array}{cc} \overline{\mathrm{W}} & -(n-u)\,\mathrm{G} \\ -(n-u)\,\mathrm{G} & \overline{\mathrm{W}} \end{array}$$

and the continuant

$$\begin{array}{ccccc} & \dfrac{-(2u+3)\,(n-u-2)}{2u+5}\,\mathrm{G} & & \dfrac{-(2u+4)\,(n-u-3)}{2u+7}\,\mathrm{G}\ldots & \\ \overline{\mathrm{W}} & & \overline{\mathrm{W}} & & \overline{\mathrm{W}} \\ & \dfrac{-1\,(n+u+1)}{2u+3}\,\mathrm{G} & & \dfrac{-2\,(n+u+3)}{2u+5}\,\mathrm{G}\ldots, & \end{array}$$

---

* *Cf.* Metzler, 'Proc. Roy. Soc., Edin.,' vol. 34, p. 223 (1914).

which is seen, by interchange of factors between conjugate elements, to be of the same type as $K_{n,u}$, being, in fact, $K_{n-1,u+1}$. Hence

$$K_{n,u} = [\overline{W}^2 - (n-u)^2 G^2] K_{n-1,u+1}.$$

This recurrence relation shows that the roots of $K_{n,u} = 0$ are $\pm (n-u) G$, $\pm (n-u-2) G$, $\pm (n-u-4) G$ ......, down to $\pm G$ or 0 according as $(n-u)$ is even or odd. Hence, using the relation between alternate sub-chains of successive chains, we see that the roots of the chain m are $\overline{W} = sG$, where s has any integral value from $-(n-m+\frac{1}{2})$ to $n-m+\frac{1}{2}$. The first sub-chain comprises the odd roots, and the second the even, or vice versa, according as $n+m-\frac{1}{2}$ is odd or even. The two central chains therefore give all the levels of the Stark effect as originally found by Schwarzschild and Epstein.

The determinant found by Schrödinger* for the levels in the Stark effect may readily be dealt with by the same method.

Summary.

The influence of an electric field on the fine-structure of the levels of the hydrogen atom is investigated on the basis of the wave-equations of Darwin and Dirac. It is found that in weak fields all the fine-structure levels except the highest split up into several, with separations proportional to the first power of the field; in the highest level the separations are proportional to the square of the field. In stronger fields the fine-structure of each of the equidistant levels of the ordinary Stark effect is found. The conclusion as regards lines is that in weak fields the fine-structure components will show a Stark effect in which some of the separations are proportional to the first power of the field, while in stronger fields each Stark component will, in general, have a simplified fine-structure. The case of H_α is worked out in detail.

I must express my thanks to Prof. Darwin for much helpful criticism and advice.

---

* 'Ann. der Physik,' vol. 80, p. 437 (1926).

Statistical Experiments on the Motion of Electrons in Gases.

By R. d'E. Atkinson, M.A.

(Communicated by F. A. Lindemann, F.R.S.—Received February 20, 1928.)

While the Quantum Theory is now almost universally accepted as the basis of discussion, when any question of energy-interchanges between molecules and free electrons is being considered, there is one collection of papers, of quite appreciable extent, that claims to treat part of the subject on totally different assumptions; these are the papers describing experiments performed in the laboratory of Prof. Townsend in Oxford.* They do not claim to account for any of the results obtained by investigators elsewhere, but they do claim to have established phenomena which it is quite impossible to fit into the otherwise universally accepted views, and which necessitate, or at least justify, a number of different assumptions which Prof. Townsend formulates. The position can only be cleared up by reconciling the experiments of the one party with the theories of the other, and it is this task which will here be attempted.

The experiments to be considered are of a quite different type from those which claimed to establish the existence of critical potentials; moreover, if the Quantum Theory statements are broadly correct, they are of a type unsuited to the problem. None the less, they would probably have led unaided, if not to the same conclusions, at least to the discovery of their own unsuitability, if it had not been for a somewhat unfortunate coincidence. This is, to put it briefly, the fact that the term-systems of the principal gases studied contain metastable states among their lowest excited levels. This fact is relatively unimportant in the critical potential work, but effects an enormous, and apparently quite unsuspected, disturbance when statistical high-pressure methods are employed, as is always the case in the work of Prof. Townsend.

The main line of experiment involved consists in studies of the rate of growth of currents in a parallel-plate apparatus as the potential is increased; a small initial current is provided by an external source such as ultra-violet light falling on one electrode, and the fields and pressures are such that the electrons produce fresh ones as a result of collisions with gas-molecules. The distance between the plates is made continuously variable, the pressure (p) and field (X) are kept constant in any one set of measurements, and the current-distance curve (which may thus also be called a current-potential curve) is then exponential for a certain range of potentials. Further, if α is the exponential constant of the curves, it is found that for any one gas α/p is a function only of X/p, as is required by the theory developed. Some subsidiary experiments, especially on the rate of rise of current at much smaller potentials (20 volts), are also taken into account, and the results are used to infer ionisation potentials and other data by a method which, though certainly indirect, probably employs the simplest hypotheses practicable.

---

* Townsend, 'Phil. Mag.,' vol. 40, p. 505 (1920), vol. 45, p. 444 (1923), and vol. 45, p. 1071 (1923); Townsend and Bailey, *ibid.*, vol. 42, p. 873 (1921), vol. 43, p. 593 (1922), vol. 44, p. 1033 (1922), and vol. 46, p. 657 (1923); Gill, *ibid.*, vol. 42, p. 852 (1921); Townsend and Ayres, *ibid.*, vol. 47, p. 401 (1924); Townsend and McCallum, *ibid.*, vol. 47, p. 737 (1924); McCallum and Focken, *ibid.*, vol. 49, p. 1309 (1925); Townsend and Focken, *ibid.*, vol. 2, p. 474 (1926).

The most important conclusion is that the ionisation potential of helium is certainly below 25 volts; indeed, results as low as 10 volts are claimed, but the subsidiary assumption employed in that argument was very directly disproved by experiment some years before it was made; these experiments seem to have escaped Prof. Townsend's notice. It should be noted that it is not claimed that 10-volt electrons can ionise helium, but rather that no electron does ionise until it has rather more than twice the amount of energy which the atom will take up, and that the atom requires 10 volts. 21 volts is thus given as the lowest energy at which ionisation actually occurs.

The assumption of the existence of metastable states, together with the secondary effects to which they may give rise, is admittedly not so "natural" as the assumptions of Prof. Townsend; indeed, if no other experiments but these were worthy of consideration, it would be quite unjustifiable. Since, further, statistical experiments are almost always more open to alternative explanations than direct ones, it is conceivable that if such an assumption had once been overlooked, the omission should remain undetected; the investigations might well lead to no hopeless contradictions among themselves. But once we attempt to extend this somewhat limited field, the case is different. Prof. Townsend has never claimed that his theory is reconcilable with any appreciable fraction of the work done in other laboratories; on the other hand, the explanation of his experiments here to be given employs assumptions whose use for this purpose is quite incidental. If they were assumptions *ad hoc*, they might be thought complicated; but all are well known and generally accepted, though they do not appear previously to have been collected together for this particular object. Thus, if they succeed in accounting for Prof. Townsend's results, they must be accepted as the only assumptions that cover the whole experimental field.

It is necessary to consider first the papers of 1924 and subsequent years. The experimental details are here irrelevant and the conclusions are in agreement with the Quantum Theory. The experiments were devoted to removing an objection which it was believed might be raised against the earlier work, namely, that the currents there measured might, to a considerable extent, be due to photoelectric effects. The conclusion finally reached, more especially in the case of helium, is that only a small fraction of any of the currents can thus be accounted for, and that the main rise of current obtained by forcing electrons through the gas is always due to the production of new ions in the gas itself. (This conclusion was already stated as an inference from the exponential nature of the curves in the parallel-plate experiments, and the inference appears to be quite justifiable, but it would seem that it was regarded as being in need of further support.) That this particular objection would have been raised seems unlikely, since the same conclusion had already been announced by Franck and Hertz,* and by Rentschler,† in 1919; but in any case the objection is one that is easily answered on theoretical grounds. For even if the collisions that result in ultra-violet radiation were in themselves much the most probable ones, the photoelectric current would still be relatively small. The reason is that the photoelectric efficiency of metals is low (generally of the order of 0·1 per cent.), and thus collisions that are evidenced only by a photoelectric effect are placed observationally at a great disadvantage. In fact, it is a perfectly general rule in all critical-potential work that, if photoelectric currents are being measured, all currents due to ionisation processes, and even the original electron-stream itself, are turned back by opposing fields so as to prevent them from reaching the collecting apparatus. This arrangement is never adopted in any of the experiments under review, and thus, if the suggestion that the results of these experiments were due to photoelectric effects has ever really been made (which, in the absence of all references, appears doubtful), it would seem in any case to be a somewhat unsatisfactory one.

The conclusion is therefore not questioned in any quarter that 20-volt electrons in helium do (in an appreciable proportion of all their collisions) produce ions in the gas itself; since, however, the ionisation potential of helium is about 24·5 volts, and the current is probably not due either to direct collisions of electrons with impurities or to cumulative effects, there is at first sight a difficulty. This, however, is resolved by a consideration of the phenomena known as "Stösse zweiter Art,"* together with the assumption that there is always a small, but still appreciable, quantity of some impurity present. This assumption cannot in fact be questioned; in 1924† it was still considered desirable to point out that the use of ebonite as an insulator inside the apparatus could cause disturbances, and the method of measuring α has always involved the presence of a large greased joint and of apparatus which (on other grounds as well) was incapable of being heated even to 100° C. Since a temperature of at least 350° is essential‡ to free the glass from adsorbed substances, and since all metals must be heated independently to a considerably higher temperature than the enclosing glass (preferably, indeed, to bright incandescence), it is improbable that the partial pressure of water and carbon dioxide was under 10^{-4} mm. in these experiments. The exact amount of impurity is quite without effect on the explanation, and 10^{-4} mm., while it is apparently 100 times less than any definite figure actually claimed, is quite possibly 100 times more than would be enough to account for the results.§ The processes then occurring are as follows:—

---

* 'Phys. Z.,' vol. 20, p. 133 (1919).

† 'Phys. Rev.,' vol. 14, p. 503 (1919).

Below 19·77 volts the electrons rebound elastically from the helium atoms, losing only such very small amounts of energy as are demanded by the principle of conservation of momentum; when they have reached the full energy corresponding to this potential, they acquire the power to lift a helium atom to the

---

[*] A "Stoss zweiter Art" may be defined as a collision between an excited (but not necessarily metastable) molecule and a normal one, such that the potential energy of the excited molecule is removed from it in whole or in part; some of the energy thus made available may be used to put the second molecule into an excited state (or, of course, to ionise it altogether), and the balance appears always (so far as is known) in the form of kinetic energy.

[†] 'Phil. Mag.,' vol. 47, p. 403 (1924).

[‡] Considerably higher temperatures are recommended by Compton and Mohler in 'Critical Potentials' (Publications of Nat. Research Council, U.S.A.). Further points, including a number of actual pressure measurements and a suggested reason for the well-known need to heat metals strongly, are now given by Campbell in 'Phil. Mag.,' vol. 2, p. 369 (1926). In passing it may be noted that the monograph of Compton and Mohler describes a great number of experiments to which it has not been thought necessary to refer in this paper, but which are, of course, in full agreement with the Quantum Theory.

[§] For an experimental proof of the extreme sensitiveness of helium to impurities, reference may be made to Horton and Bailey, 'Phil. Mag.,' vol. 40, p. 440 (1920). It is there shown that the detection of the 20-volt break depends upon the presence of impurities which can be removed by repeated circulation through cooled charcoal, but which rapidly reappear, *even in a baked-out apparatus*, if the flow is interrupted. Similar effects were recorded slightly earlier by Franck and Knipping, 'Z. f. Physik,' vol. 1, p. 320 (1920). *Cf.* also Rentschler, *loc. cit.*

metastable 2^3S* state on collision, and after, probably, no very large number of collisions they will experience one in which they do this, losing all their energy again. Such an excited atom is large,† electrically neutral and it has usually a long life; it will thus remain nearly in one place, and in the course of, say, 10^5 collisions with other atoms will probably encounter a molecule of some impurity; the fact that it has an abnormally large electric moment may even give it a slight tendency to attract such molecules selectively. We shall assume for the present that there is no effective quantity of neon in the helium, *i.e.*, that the proportion of neon is certainly less than 1 part in 10^6; the ionisation potentials of all other likely impurities are well under 20 volts, and thus, if the encounter (as is believed to be fairly probable) does take the form of a "Stoss zweiter Art," the *molecule* will be ionised. At high pressures it is usually only the first critical potential (together with any others that may happen to lie close to it) that can be operative, and there will thus be a very strong tendency for the first critical potential in helium (which, it may be remarked, is not under any circumstances in itself a radiation potential)‡ to appear as an ionisation potential; in fact, Prof. Townsend now describes 21 volts as the lowest potential at which ionisation occurs, and a slightly lower figure was originally given by Franck and Hertz, Benade, Benade and Compton, Pawlow, Bazzoni,§ and a large proportion of those physicists who quoted their results up to the end of 1919. Somewhat similar results may be expected with neon, but the effect of radiation will be larger, since several unstable and metastable potentials here lie very close together.

It is important that the life of the metastable atom is long; the process is thus practically independent of the amount of impurity present, within somewhat wide limits. The claim* that the results of these experiments could be due to impurities only if the amount of impurity were always proportional to the amount of helium presumably holds for direct effects, but is plainly not proof against the more formidable attack based on this indirect effect, since multiplying or dividing the amount of impurity by 10 would now have very little effect on the current observed, and none on the potential at which it began to appear.

---

* For the notation, see Russell and Saunders, 'Astroph. J.,' vol. 61, p. 66 (1925). The assignation of a triplet-multiplicity to the orthohelium terms, which had been current for some time on theoretical grounds, has also been verified experimentally (Hansen, 'Nature,' vol. 119, p. 237 (1927)). The S terms themselves are, of course, single in either case. For the actual existence of the metastable state, the references are too numerous to quote, but perhaps mention may be made of the paper of Franck and Reiche, 'Z. f. Physik,' vol. 1, p. 154 (1920).

† This appears not to be (otherwise) directly verified for helium, but it has been observed in mercury (Cario, 'Z. f. Physik,' vol. 10, p. 185 (1922)), and is on the older "atomic model" views inevitable in all cases; the area of the atom should vary roughly as the fourth power of the effective quantum number.

‡ Sommerfeld, 3rd ed., p. 396, etc. Any radiation which does experimentally appear at this potential is, according to the accepted view, necessarily due to some secondary effect, and, in fact, a line at 624 Å is not observed.

§ Franck and Hertz 'V. d. D. P. H.,' vol. 15, p. 34 (1913), etc.; Benade, 'Phys. Rev.,' vol. 10, p. 77 (1917); Benade and Compton, 'Phys. Rev.,' vol. 11, p. 184 (1918); Pawlow, 'Roy. Soc. Proc.,' A, vol. 90, p. 398 (1914); Bazzoni, 'Phil. Mag.,' vol. 32, p. 566 (1916).

Moreover, the process, though essentially complex, will appear experimentally as a single one ; the effect will not depend on the square of the current, and the ion will be produced very nearly at the same time, and nearly at the same place, as that at which the collision occurred. The most important difference between this explanation and Prof. Townsend's is that, though the electron lost its energy in a collision with a helium atom, the resulting ion is not a helium ion. This possibility appears not to have been contemplated, and it seems to be one that is not very readily open to investigation by statistical methods ; if proof were required, however, Smyth's† arrangement would presumably show at once that the ion had a mass several times that of a helium atom. Such an enquiry would thus be a possible method of studying the impurities present in fairly well-evacuated vessels.

The presence of neon as an impurity would perhaps decrease the number of "Stösse zweiter Art" that resulted in a new ion, since the neon itself cannot be ionised, and the energy which it did take over would in many cases then be radiated. Neon has, however, also metastable terms to which it could fall by a radiational process, and such atoms would still contain enough energy to ionise many impurities on collision. A further property of neon is, however, important, and especially so when this gas, and not helium, is the one that is being ostensibly studied. In those cases in which the atom does return to its normal state by a radiational process, the radiation will be of large enough frequency (about 16 volts at the least) to ionise many molecules by simple absorption. The efficiency of this process is, under favourable circumstances, probably high, and the radiation may in addition be reflected from the electrodes and glass if it has not been at first absorbed ; the ionisation can clearly occur at any point, not merely in the electron-stream, and it will thus appear as though the electron-stream, or the positive ion-stream, was scattering more than was to be expected. This result has been obtained by Prof. Townsend, but it is simply given as an isolated datum, and no explanation has been suggested for it.

---

* 'Phil. Mag.,' vol. 46, p. 658 (1923).

† 'Nature,' vol. 111, p. 810 (1923).

The possibility of producing ions well below the ionisation-potential of the helium itself being thus established, it remains to account for the fact that above about 60 volts the current rises exponentially. At high pressures, such as are always employed in these experiments, it is probable, as we have seen, that very few electrons ever reach an energy much above that of the first critical potential, and if the amount of impurity present is sufficiently small, they have, apart from the small elastic losses due to transfer of momentum to the molecules in elastic impacts, no way of losing energy below that figure. The regions where increase of current occurs should thus at first sight be layers parallel to the plates, spaced apart by distances corresponding to 20 volts + the sum of the "elastic" losses; and increasing the distance apart of the plates should result in current-increases by steps, as these distances were successively reached, and not in an exponential curve. It is evidently necessary, therefore, to find some means of destroying this simple nature of the electron-stream.

A small, though not negligible, factor tending to complicate the stream is provided by the occasional collisions of an electron with a molecule of some impurity directly. For even if such collisions only occur in 10^{-4} of the cases, the probability that one will occur somewhere in any given rise of any given electron from 0 to 20 volts is quite appreciable; and practically in every such case the electron will lose all its energy.

A probably more important factor is provided by the normal diffusion of the metastable atoms between the time of their formation and the "Stoss zweiter Art." The distance they will diffuse cannot be given *a priori*, since it depends both on the amount of impurity present and on the diameter of the metastable atom, both of which are unknown. But it is almost certain that the metastable atom will be very much larger than a normal one, and application of the formula $dR^2/dt = 2lV$ indicates that it may well diffuse some millimetres, *i.e.*, several volts up or down the field, before it gives up its potential energy. Indeed, the difficulty is rather to show that it will not actually diffuse too far than vice versa. Since it will neither gain nor lose energy during the diffusion, the new electron ultimately produced will have an energy, at any given point in the field, some volts larger or smaller than that of the original electrons at that point. It is easily seen that this branching will have made the stream very complicated even by the time 45 volts have been reached, and there appears no reason to doubt that by about 60 volts the confusion (due to this cause alone) should be fully adequate to account for the observations. It may also be remarked that probably the rate of growth of the current in its earliest stages is limited primarily by diffusion of the metastable atoms to the electrodes, and is thus almost entirely without significance. This possibility entirely undermines the argument by which ionisation potentials of 10 volts were obtained; arguments based only on the values of α would, however, be unaffected, since α does not exist below about 40 or 50 volts.

The above considerations are scarcely capable of exact quantitative amplification, since the number of unknowns introduced by the statistical method is so large; but qualitatively they appear very well suited to account for the existence of an exponential curve, starting at about the right potential. If this view is then correct, it is evident that the exact figure for the exponential constant obtained is almost valueless, since it depends on so many unknown interacting factors, such as the nature and relative proportions of the various impurities present, the amount (if any) of energy which the newly formed electron may possess, the relative effective diameters of the molecules in question, and the absolute efficiency of the "Stösse zweiter Art" to produce an electron at all. We may, however, neglect most of these complications (some of which tend, in fact, to cancel out against each other) for the sake of showing that at any rate the order of magnitude of the values of α is about what would be expected.

For the case $X/p = 5$, $\alpha/p = 0 \cdot 12$; if, for simplicity, we take $p = 1$, it follows that after travelling $0 \cdot 693/0 \cdot 12 = 5 \cdot 78$ cms. each electron produces a new one. In the course of this distance it has fallen through $28 \cdot 9$ volts. The inelastic collision absorbs $19 \cdot 8$ volts, leaving $9 \cdot 1$ volts for the loss in the large number of elastic ones. The number of collisions in a distance a is $3a^2/4\lambda^2$ (where λ is the mean free path) according to Hertz, but this result is obtained on the assumption that the elastic losses are negligible. The figure will obviously be greatly reduced when they are so numerous that a third of the electron's energy has been dissipated by them, because the velocity of drift is inversely proportional to the velocity of agitation; since thus the actual velocity has been reduced to 4/5 by the elastic losses and the velocity in the direction of the field has been increased to 5/4, it seems not unreasonable to put the number of collisions at about 2/3 of that given by Hertz's formula. The loss on each collision is $2 \cdot 73 \, . \, 10^{-4}$ of the energy possessed at the moment, so that for n collisions the loss is $2 \cdot 73 \, . \, 10^{-4} \, . \, nE$ if E is the average energy of the electron, the average being taken over the number of collisions, not over the time or the straight-line distance travelled. E will be larger than $19 \cdot 77/2$, because more collisions occur at the higher energies; on the other hand, the mean free path increases steadily with the voltage between 4 and 20 volts, and this will have a compensating effect;

without going into details, which would involve a somewhat complicated reckoning, we may plausibly put E at about 10 volts, so that the total elastic loss comes out at

$$\tfrac{4}{3} . 2\cdot73 . 10^{-4} . \frac{3 . 5\cdot78^2 . 10}{4 . 6\cdot4^2 . 10^{-4}}$$

if we take the figure of Townsend and Bailey for the mean free path; and this works out to 11·1 volts. The error of 2 volts (in 30 volts) is probably adequately accounted for by the facts that on the mathematical side several quantities have been averaged which should have been treated in detail, and that on the physical side it has been assumed that no part whatever of the increase in current is due to collisions of electrons with impurities (*including neon*). Until these points are disposed of, the values of α thus constitute no challenge to the Quantum Theory.

It is easily seen that doubling the pressure and field, and halving the distance in which an ion is produced, will leave the figures otherwise unaltered, so that α/p is, as before, a single-valued function of X/p. It is, further, to a first approximation, proportional to X/p, but falls below the straight-line figure for small values of X/p (owing to the elastic losses), and perhaps also for large ones (owing to diffusion); this is actually the form of curve experimentally obtained.*

It is possible to obtain a somewhat similar verification if we accept the statistical figure for the "average loss of energy on collision," but this conception appears so very questionable that the other method has been preferred. The "average loss of energy" was, namely, first considered for the case where no ionisation is taking place; it was there expressed as a fraction, not of the energy possessed by the electron at the time, but of *its* average energy†; on removing the restriction it was not apparently defined explicitly whether this average is taken over the life-history of one electron, in which case it might be with respect to time, to straight-line distance travelled, or to the number of collisions, or whether it is taken over all electrons. For the case of no ionisation, many of these possibilities coincide, but as soon as ionisation sets in, there are very large differences between the energies of different electrons, which alone would seem to make any process of averaging most undesirable. In addition, there are now the new slow electrons to consider, and these invalidate the formula from which the "average loss" is obtained. This last point has apparently been recognised in the most recent work,‡ but the full nature of the difficulties introduced by this extension of statistical methods appears still not to have been realised.

---

* Townsend and Ayres, 'Phil. Mag.,' vol. 47, p. 410 (1924).

† Townsend and Bailey, 'Phil. Mag.,' vol. 42, p. 888 (1921), and vol. 44, p. 1045 (1922).

‡ Townsend and Focken, 'Phil. Mag.,' vol. 2, p. 481 (1926).

The method of obtaining the " average loss " appears to contain a very grave error somewhere, even if it is restricted to the case of small fields and no ionisation. For the figures given,* both for argon and helium, in this case imply that the average loss is only about half what it must be if the atoms took up no internal energy from the electrons at all; there is then only the small loss $2m/M$ (m being the mass of the electron and M that of the atom) due to the inevitable transfer of momentum to consider. There appears no way at all of accounting for a smaller loss than this, whatever the process of averaging; but on two separate occasions results much below this figure have been published without comment. Even when the departure is not as serious as 50 per cent., it is often very appreciable. For example, there is a considerable range of fields for which the loss in helium† is given as $2 \cdot 5 \,.\, 10^{-4}$, which is 10 per cent. smaller than it possibly can be, and this figure is also quoted‡ in a subsequent paper instead of the $2 \cdot 73 \,.\, 10^{-4}$ which it clearly indicates as the correct one. The exact cause of the error need not be enquired into here, but the fact that such a serious error has been overlooked cannot but create a suspicion that in more obscure cases also, such as the calculations of mean free paths, agitation energies and so on, equally essential considerations may have been omitted.

At this point we may consider what distribution of velocities must be present in the stream if the current-curve is, in fact, to be exponential. Once the distribution has reached its most probable form (a necessary condition for the elimination of all steps in the curves), it is, in fact, fully determinate; the deciding factors are not at all of the type which ordinarily lies behind a Maxwellian distribution, and none of them is at our disposal for the purpose of further developing a theory, since they are themselves the atomic characteristics which it is hoped to infer. For example, if the electrons do always lose all their energy at about 20 volts, and if there were no diffusion of the metastable atoms, the final distribution (which could not, however, be easily reached in that case) would clearly be a simple exponential one. For the rate of production of new ions at any point A is proportional to the number of 20-volt electrons passing A; the larger rate of production at a point, say, 2 volts down the field (later) is similarly proportional to the number of 18-volt electrons passing A, and so on

---

* For helium, 'Phil. Mag.,' vol. 46, p. 661 (1923); for argon, *ibid.*, vol. 44, p. 1045 (1922).

† Townsend and Bailey, 'Phil. Mag.,' vol. 46, p. 662 (1923).

‡ Townsend and Focken, 'Phil. Mag.,' vol. 2, p. 480 (1926).

(elastic losses being neglected), since all these 18-volt electrons will by then have attained 20 volts, and no other source of 20-volt electrons exists.* If the range over which loss of energy is possible were not the very restricted one here considered, but the very wide one required by Prof. Townsend's theory, the distribution would depart from the exponential form, perhaps very greatly; and if, as is here argued, the new electrons are not produced at the points where the original collisions occurred, a further complication is introduced; but in every case the distribution is, in principle, calculable whatever theory we may adopt, and it is not permissible, as has often been done,† to *assume* a distribution (whether Maxwellian or otherwise), at all. Further, the distribution must extend fully over a range equal to the usual loss of energy in an inelastic collision, *i.e.*, even on Prof. Townsend's view at least 10 volts, and there must as a rule be an excess of the slower electrons in the range, since these are the ones that will be responsible for the large increase of current at a late point in the field. And since each electron goes through a cyclic determinate process, independently of the others, in which its energy rises slowly and falls suddenly, the conditions are so different from the case where no ionisation is taking place, and where *each* electron does really reach a steady state, that it seems very doubtful whether the practice of discussing the "average energy of the electrons" can be anything but confusing.

The statistical interpretation of these experiments is thus open to distrust on its own merits alone; if, however, it be granted that the alternative interpretation that has here been given is an admissible one, it follows in any case that the statistical one is not an admissible one at all. For it conflicts with so many experiments which cannot be questioned that its only justification was that it was still necessary to account for some results. It is also clear that the statistical experiments, interpreted in the new way, are unsuited to give any estimate of the critical potential of helium; the probable error in the original interpretation of them has never been claimed to be under 1 volt, and the Quantum Theory view cannot really lead to any definite figure at all. The more recent measurements by the direct methods give results which are often only a few hundredths of a volt in error, the optical confirmation being, moreover, in some cases entirely subsequent.

---

* The distribution here considered is in terms of the number *crossing a given plane* per second. The distribution in terms of a length dx of the field is, of course, obtained from this by dividing the number in each energy-range by the velocity drift appropriate to that energy. The form here given is simpler.

† Townsend, 'Phil. Mag.,' vol. 40, p. 509 (1920); Townsend and Bailey, 'Phil. Mag.,' vol. 42, p. 891 (1921), and vol. 44, p. 1044 (1922).

It remains to criticise very briefly the evidence provided by experiments of the direct type, and the objections raised against them in the papers under discussion. On careful consideration, one cannot escape the conclusion that these objections are based on misapprehensions as to the actual claims of the Quantum Theory. For example, the statement is made* that, according to the Quantum Theory, collisions below the first critical potential are assumed to occur without any loss of energy whatsoever; the theory is accordingly criticised on the grounds that this would violate the principle of the conservation of momentum, that it cannot be substantiated from any direct measurements of energy-losses, and that it is at variance with the results of the statistical experiments. In fact, a considerable portion of the attack in the latest of these papers is based solely on the statement that according to the Quantum Theory this remarkable type of impact, involving no loss at all, occurs at small potentials. The statement cannot be justified. The phrase "collisions without loss of energy" is occasionally found in subsequent references to some experiments† which claimed to show, *with an accuracy then estimated at from* 1/3 *to* 1/10 *volt*, that at low potentials "no loss" occurred on reflexion; the loss involved in an elastic collision, however, is of the order of 1/1000 volt, so that it is immediately evident these experiments were concerned only with possible inelastic losses. But in case any doubt might arise, this was very early pointed out; in one of the first theoretical papers‡ on the subject, Franck and Hertz enquired what phenomena were to be expected if the electron lost no appreciable energy below a certain potential, and then lost it all and produced a fresh electron by collision; in the last paragraph they explained that the theory was only rough, and that they were not concerned with the small difference between "no loss" and "elastic collisions," the position on the experimental side being, to quote their words, "if the collisions were strictly elastic, the electron, as is easy to see, would lose in helium, for example, roughly 1/4000 of its energy at each encounter; *such a loss cannot be established by our methods.*" The phrase "collisions without loss of energy" was thus merely a loose alternative for the much more usual "elastic collisions§" in cases where the two were experimentally indistinguishable; originally it was always an abbreviation for "collisions without the Townsend loss of energy," in reference to a now discarded suggestion of Prof. Townsend[*] himself for deducing ionisation potentials from the constants of his well-known "collision formula." The phrase would perhaps never have been used at all if the main purpose[†] of the above experiments had not been to disprove this suggestion.

---

* 'Phil. Mag.,' vol. 2, p. 475 (1926).

† *E.g.*, Franck and Hertz, 'Verh. deut. Phys. Ges.,' vol. 15, p. 380 (1913), and 'Phys. Z.,' vol. 17, p. 430 (1916).

‡ 'Verh. deut. Phys. Ges.,' vol. 16, p. 12 (1914).

§ Franck and Hertz, 'Verh. deut. Phys. Ges.,' vol. 15, pp. 384, 618 (1913), *ibid.*, vol. 16, pp. 12, 467 (1914), 'Phys. Z.,' vol. 17, p. 434 (1916), *ibid.*, vol. 20, p. 132 (1919); Franck and Knipping, 'Phys. Z.,' vol. 20, p. 481 (1919); Compton, 'Phys. Rev.,' vol. 7, p. 509 (1916); Tate, *ibid.*, p. 686; Townsend, 'Phil. Mag.,' vol. 40, p. 506 (1920); Benade, *loc. cit.*; Benade and Compton, *loc. cit.*; Goucher, 'Phys. Rev.,' vol. 8, p. 562 (1916); Davis and Goucher, 'Phys. Rev.,' vol. 10, p. 101 (1917), etc.

The elastic loss was, however, very thoroughly investigated shortly afterwords, on both sides of the Atlantic. Hertz, on the one hand, and Benade and Compton on the other, working independently, established the facts that (to take their conclusions as a whole) at all actual (not average) energies below about 20 volts in helium the collisions were strictly elastic, the observed loss per collision being in good agreement with the calculated value,[‡] and that at 20 volts something which appeared to be ionisation of the gas set in. The papers form altogether a very extensive study of this question,[§] and would not seem to leave room for any misunderstanding. That genuine violations of the principle of the conservation of momentum do occur in the statistical work has already been shown.

In general the criticisms offered display misapprehensions comparable with the above, and in some instances the critical-potential work is mentioned only to be dismissed. One point, however, may be mentioned in detail. The "method of Franck and Hertz[||] (method of inelastic collisions) provides definite evidence as to the energy possessed by the colliding electron (and by the new one also produced, if there is one) immediately after a "critical" collision; Prof. Townsend states categorically that there is no evidence on the question,[¶] and his own theory is completely incapable of accounting for the abrupt decrease in current with increasing potential obtained in these experiments, even if that did not occur at predicted points. He has apparently failed* to distinguish them from the work by "Lenard's method," which, though superficially similar, in that the same apparatus can be used for both, is directed to a different aim and employs a quite different principle.

---

[*] 'The Theory of Ionisation of Gases by Collision,' London, 1910.

[†] Franck and Hertz, 'Verh. deut. Phys. Ges.,' vol. 15, p. 930 (1913); and elsewhere.

[‡] The calculated figure $2 \cdot 9 . 10^{-4}$, given by Hertz is reduced to $2 \cdot 73 . 10^{-4}$ by insertion of more recent values for the masses; the experimental figure was $2 \cdot 7 . 10^{-4}$. Benade and Compton showed that the elasticity could not differ from unity by more than 10^{-4}. Both these conclusions appear to need some modification in the light of Hertz's later considerations, but this is obviously without effect on the point at issue.

[§] Benade, *loc. cit.*; Benade and Compton, *loc. cit.*; Compton and Benade, 'Phys. Rev.,' vol. 11, p. 234 (1918); Compton, 'Phys. Rev.;' vol. 21, p. 717 (1923), and vol. 22, pp. 333, 432 (1923); Hertz, 'Verh. deut. Phys. Ges.,' vol. 19, p. 268 (1917), and 'Z. f. Physik,' vol. 32, p. 298 (1925).

[||] See especially Franck and Einsporn, 'Phys. Z.,' vol. 22, p. 390 (1921). The earlier examples, however, which extend over a wider range of potentials, are also very convincing.

[¶] 'Phil. Mag.,' vol. 45, p. 1079 (1923).

This paper is not a defence of the Quantum Theory as a whole, and the arguments have been almost entirely confined to the one branch of the subject, namely, the purely electrical, in which opposition is still offered; but the real strength of this aspect of the Theory lies in the remarkable unification of all branches of physics which *it* effects, a unification which supplies the investigator with an extraordinary array of new problems, and, at the same time, with powerful new weapons that almost invariably stand the test of use. Mention may perhaps be made of the work of Hertz in 1923† (almost all earlier work on this point is suspect) in which the (electrical) energies necessary to excite spectral lines to a (spectrographically) observable intensity were at last shown to be those required by the Theory, and of the help which may therefore legitimately be expected, and which is, in fact, obtained, from direct energy-measurements in analysing complex spectra, whether by temperature-classification or as a result of actual critical-potential measurements. It is true that the mere attractiveness of this unification won for it most of its best-known supporters many years before the evidence itself was in any way complete; the fact, however, that they jumped on comparatively slight grounds to their conclusions cannot, now that these conclusions have been at every point so rigorously verified, be urged in disparagement of their judgment.

[*Added, April* 22, 1928.—In a paper of April, 1928, Professor Townsend and Dr. McCallum have shown that extra precautions in purification make a considerable difference in the figures obtained. The new results in no way invalidate the contentions set forth above.]

---

* 'Phil. Mag.,' vol. 47, p. 737 (1924), vol. 45, p. 1071 (1923).

† 'Z. f. Physik,' vol. 22, p. 18 (1924).

Observations on the Band Spectra of Mercury.

By Lord Rayleigh, F.R.S.

(Received April 24, 1928.)

(Plate 4.)

This paper records some observed facts on the elusive subject of mercury band spectra. I must confess at the outset that I have little that is definite to offer in the way of interpretation.

§ 1. *Excitation of Mercury Band Spectrum by the Continuous Spectrum of Hydrogen.*

It was noticed in absorption experiments on mercury vapour at atmospheric and higher pressures* that the well-known green mercury fluorescence was quite marked where the light of the continuous hydrogen source entered the vapour. The fluorescence has hitherto been chiefly studied by using the mercury arc or the aluminium spark as a source. It appeared worth while to investigate the fluorescence spectrum with this new means of excitation.

It was found that the visual fluorescence was nearly extinguished by a filter of 2 mm. of calcite, or by 9 mm. of fused silica. The calcite filter cuts off the spectrum at about λ 2120. It appears, therefore, that the relevant region of absorption is that which is believed to begin at the resonance line 1850, extending towards longer waves. Absorption in the region from 2345 towards shorter waves, or in the region 2537 towards longer waves, is not concerned to an important extent in producing at any rate the visual fluorescence. Whether these absorptions contribute to the ultra-violet fluorescence is an important question, which has not yet been investigated with the hydrogen source.

The bright visual fluorescence extends, about 1 to 1·5 cm., into the vapour at atmospheric pressure, which is about the thickness of vapour required to absorb the radiations which are stopped by the 2 mm. calcite filter ; a fainter green fluorescence extends beyond this, and is doubtless due to excitation by longer and more penetrating waves. As with other modes of excitation, the fluorescence is not seen unless air is completely expelled by the vapour of mercury.

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 702 (1927).

In these preliminary experiments the vapour was at atmospheric pressure, but in photographing the spectrum it was found better to use a lower pressure (6 cm.) which allowed the brightest part of the fluorescence to spread well away from the wall of the containing vessel, for avoidance of false light. The containing vessel was a tube 28 cm. long and 2·5 cm. in diameter laid alongside the hydrogen discharge tube, so that the light entered by the side. It had a flat end plate fused on, through which the fluorescent spectrum was photographed. The far end of the tube was drawn down and bent aside in the usual way to afford a black background. The bottom of the tube was covered with a layer of mercury which was kept boiling under a reflux condenser; the top end of the condensing tube was connected to a large bulb, which contained air at a reduced pressure, in this case 6 cm. The mercury was boiled under this pressure by resistance heaters. All air was thereby pushed out of the optical tube up to the point in the condensing tube where the mercury liquefied.

A small Hilger quartz spectrograph was used. The fluorescence spectrum (Plate 4, reproduction A) was peculiar, and differed in several respects from that obtained by other modes of excitation. The usual broad structureless bands in the visual and ultra-violet regions occur in about equal photographic intensities, and so far there is nothing exceptional. But when we come to the region of the resonance line 2537 a most unusual distribution of intensities is observed. The narrow band 2540 is in fact the only feature in this neighbourhood. The resonance line 2537 itself is absent, and so is the stretch of continuous spectrum which reaches from this point with diminishing intensity towards the red. The maximum at about 2650 to which Houtermanns* has drawn attention is likewise absent. It is, of course, possible that any or all of these features might come up with more intensity in the source or longer exposure. But, so far as I am aware, the band 2540 has never been obtained in emission before without the resonance line accompanying it in much greater intensity. A spectrogram with the mercury lines in comparison with the fluorescence spectrum shows quite definitely the absence of the resonance line and presence of the band.†

In an earlier investigation‡ the spectrum of mercury vapour as excited by an aluminium spark is shown. If the vapour is fairly dense, then, as originally noticed by R. W. Wood, the resonance line does not appear, and the band spectrum is conspicuous. But the case is essentially different from that of hydrogen tube excitation. For with aluminium excitation the band 2540 is dark on the bright background of continuous spectrum which reaches from a little below the resonance line towards long waves, whereas with continuous hydrogen spectrum excitation 2540 is bright on a dark background. Reference should be made to the plate accompanying the paper cited.

---

* 'Z. f. Physik,' vol. 41, p. 140 (1927).

† This particular plate, with the mercury comparison spectrum, was not used for the reproduction, as it was scarcely intense enough for the purpose.

‡ 'Roy. Soc. Proc.,' A, vol. 111, p. 456 (1926).

We next come to the features of the spectrum at 2345. Here, again, hydrogen excitation gives a peculiar result, for of the series of bands 2345, 2338, 2334, etc., we find only the first. When these bands are observed in absorption, or when excitation is by electric discharge, the other bands of the series are observed as well. (*See* reproduction B and D on plate 4.*) It is true that the first band is always the most intense; and the absence of the others may perhaps be thus accounted for. But, on the short wave side of the bands there is a region of continuous emission, and in the present spectrum this has considerable intensity over a wide range of wave-lengths. Its intensity relative to the first band is much greater than in the electrical method of excitation. Compare plate reproduction, A hydrogen excitation, B electrical excitation. It appears clear from the present result with hydrogen tube excitation that this stretch of continuous spectrum cannot be regarded as capable of resolution into a continuation of the series of bands. For in this case it sets in immediately after the first band, the other bands being absent.

In a former investigation† I found that this continuous spectrum, as observed in the excited vapour (after-glow) from the discharge, first became perceptible at the forbidden line $1\,{}^1S_0 - 1\,{}^3P_2$ (2270), extending from this line with increasing intensity towards long waves, when it was merged in the series of bands 2345, etc. In the present instance the forbidden line is not present, and this reach of continuous spectrum is traceable as far as 2150, thus extending much beyond the position of the forbidden line. It does not follow from this that the interpretation put on the former observations was wrong, for the conditions were different, the excitation of this part of the band spectrum not improbably depending on atoms in the ${}^3P_2$ metastable state; as evidenced by the observed occurrence of the forbidden atomic line. In the present case, where the forbidden line does not occur, there is not the same reason for attaching special importance to this position in the spectrum, as likely to represent the highest excitation concerned.

A similar long extension of the continuous region far past λ 2270 has been observed in the electric discharge itself under certain conditions, *see* reproduction C on the plate, which repeats B but with less density of printing so as to show the rather faint continuous spectrum, though with unavoidable loss of the detail of the bands 2345, 2338, etc. So far as the available evidence goes, it is only in the excited vapour or after-glow that the limitation to wave-lengths longer than 2270 holds good. When the vapour is actually under stimulation by the discharge or by the hydrogen tube, the reach of continuous spectrum extends to considerably shorter wave-lengths.

---

* For absorption, see also 'Roy. Soc. Proc.,' A, vol. 116, p. 718 (Plate 22) (1927).

† 'Roy. Soc. Proc.,' A, vol. 114, p. 638 (1927).

§ 2. *The Series of Bands* 2345, 2338, *etc., under High Dispersion.*

It was found in the investigation already cited* that these bands as observed in absorption with increasing quantities of mercury were underlaid by a finer structure of bands degraded to the red spaced at about 1 Å. It was thought on the whole that this finer structure was an essential part of the phenomenon, difficult of observation indeed, but not to be separated from the coarser series of several Ångströms spacing which had been known long before. I now find that this view needs revision.

It was always intended to examine the bands in emission with large resolving power, the only difficulty being to get a bright enough source. The source ultimately adopted was a high tension alternating current discharge in dense vapour, between mercury electrodes. The discharge tube was 9 mm. diameter and 6 cm. long. It was viewed end on. By means of heating coils the mercury in the end vertical portions was kept boiling under an atmosphere of 6 mm. pressure with reflux condensation, in the way already described above (p. 350). An alternating current of about 100 milliamperes was passed. Under these conditions the band spectrum was very bright though, of course, by no means free from line spectrum.

Photographed with the large Hilger Littrow spectrograph, this discharge showed the bands in question sufficiently well. See reproduction D on Plate 4. Four of them are visible, and there is *no trace whatever* of the underlying bands of 1 Å. spacing, which in absorption we see so clearly with this instrument. The conclusion would seem to be that the fine structure (1 Å. spacing) and the coarser structure (4 to 7 Å. spacing) are not merely different aspects of one complex phenomenon, but represent two superposed phenomena. In absorption they occur together. In emission the coarser structure is obtained alone.

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 718 (1927).

§ 3. *Bands* 2482 *to* 2476 *and* 2476 *to* 2410 *under High Dispersion.*

In examining spectrograms of the high tension discharge in fairly dense mercury vapour, another band structure was noticed, which, it is believed, has not been described hitherto. There are two emission bands mentioned by Stark and Wendt[*] who give the wave-length ranges as 2482 to 2476 and 2477 to 2470, the former being the more intense. These are shown with the adjacent mercury lines on Plate 4, reproduction F, with my own measurements of the heads, which, as will be seen, are fairly well defined in position. Stark and Wendt describe these bands as continuous. They are now photographed on a larger scale than previously, and it appears that the brighter of the two is not continuous, but exhibits a remarkable structure. See reproduction E on the plate.[†] The structure, it will be noticed, consists of alternate maxima and minima, the intervals tending to converge towards the head. The position of these maxima, and of the heads of the two bands, were measured on several negatives relative to the atomic mercury lines and to an iron comparison spectrum. The mean results were as follows :—

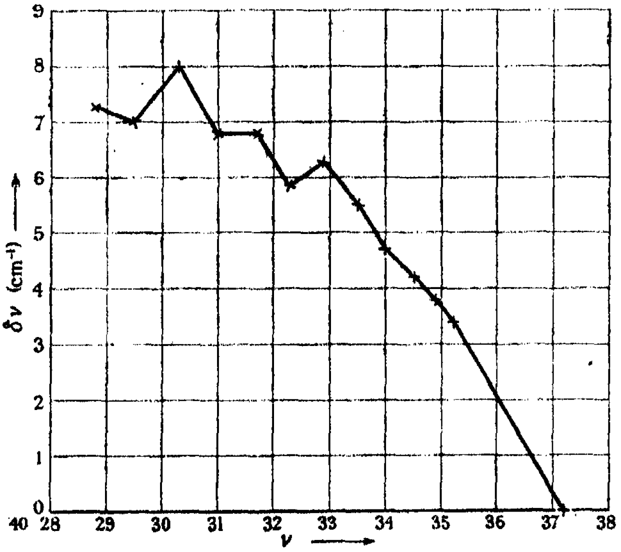
Description.	Wave-length.	Wave-number.	Difference.
Maxima in strong band	2481·60	40284·4	
	2481·15	40291·7	7·3
	2480·72	40298·7	7·0
	2480·23	40306·7	8·0
	2479·81	40313·5	6·8
	2479·39	40320·3	6·8
	2479·04	40326·1	5·8
	2478·65	40332·4	6·3
	2478·31	40337·9	5·5
	2478·02	40342·6	4·7
	2477·76	40346·8	4·2
	2477·53	40350·6	3·8
	2477·32	40354·0	3·4
Head of strong band	2476·20	40372·3	
Head of weak band	2469·70	40478·5	

---

[*] 'Phys. Zeits.,' vol. 14, p. 563 (1913).

[†] The same structure is seen on the original of reproduction F, but in order to bring up the fainter band, it was lightly printed, so that little of the structure in the stronger band can be seen.

It will be noticed that the interval between the third and fourth maximum is larger than any other interval. This excess was independently confirmed on three different negatives, and appears to be real, indicating a definite lack of regularity in the series. The last six of the intervals, however, diminish pretty regularly tending to converge to zero at a point not very far from the position of the head. This is indicated by the graph (see figure) which shows the mean



wave-number as abscissa and the interval in wave-numbers as ordinate, the head being treated as the position of zero interval. There is no attempt to draw a smooth curve, the successive points being joined by straight lines.

The second and fainter band reaching from about λ 2476 to λ 2470, when examined under small dispersion looks exactly similar to the other in its general aspect. It is surprising, therefore, to find that it shows no structure under the higher resolving power of the Littrow spectrograph. This is certainly not due to lack of intensity, as I have verified by examining weaker photographs of the stronger band, which show the structure clearly enough. The two bands are almost certainly related intimately in some way, as they appear under similar conditions, and (apparently) in constant relative intensity, so far as qualitative examination can show. I have not met with anything at all analogous in the literature of band spectra.

It is important to determine whether these bands occur in absorption. This point can be tested advantageously because they are situated in a part of the spectrum which is not readily lost in the general opacity which covers other regions as the amount of vapour, is increased. I have used a column of saturated mercury vapour, 87 cm. long at 190 cm. pressure, but no trace of these bands was to be seen in absorption. There would be no difficulty in carrying the test a good deal further if it were judged worth while.

It appears then that the bands in question do not result from a transition to the unexcited electronic state of the molecule.

Another point of interest is that these bands occur in the excited stream of vapour (after-glow) rising from a low tension discharge with hot wire cathode.* In photographs taken with an exciting discharge current of 3·5 milliamperes and 12 hours or more exposure the stronger band was well seen. The weaker one was suspected, but, having regard to relative intensities, it could not be expected to show up strongly.

As already remarked, the relation between the stronger band with structure, and the adjacent weaker one with apparently no structure, is not easily paralleled. Suppose, however, that we ignore for a moment the mysterious relation between these two bands, and enquire whether we can regard the structure of the stronger band as either vibrational or rotational, in accordance with the quantum theory of band spectra. Either view meets with difficulties.

The close spacing of individual maxima is more in accordance with a rotational origin, and, as my friend Mr. W. Jevons has pointed out to me, the presence of a head to the series decidedly favours the same alternative; for known series of bands forming a vibrational progression invariably tend to die out towards the point of convergence, the position of which can only be assigned by extrapolation. The same applies to a vibrational sequence.

On the other hand the general appearance of the structure is not very like that of recognised rotational ones. We should not expect to be able to resolve the lines of a mercury band in the neighbourhood of the band origin, on account of the large moment of inertia. But, even at a distance from the head the present structure has little of the appearance of *e.g.*, the resolved cyanogen bands, which consist of very fine lines separated by considerable intervals. The individual maxima in reproduction E are certainly very far from being fine lines of this kind. It may be thought that this is due to deficient resolving power, but this idea can hardly be sustained. The atomic mercury lines in E are broad, it is true, but this is merely because they are over exposed. The iron comparison lines which were put on several of the photographs were very sharp in comparison with *e.g.*, the maximum at λ 2481·6. It may be possible to get over this objection by assuming special distributions of unresolved lines; but upon the whole the difficulty stands out that a rotational structure is in general repeated along the spectrum in virtue of the fact that a generally similar series of rotational levels is attached to every vibrational level. So far as I have been able to observe the present structure is not repeated, but is unique.

---

* The arrangement is described in 'Roy. Soc. Proc.,' A, vol. 114, p. 620 (1927).

If we turn to the alternative of a vibrational structure, we are met with the difficulty that recognised vibrational progressions are much more widely spaced than the series under consideration, the initial interval is only 7 cm.$^{-1}$. In the list of constants of band spectra given by Birge* the closest vibrational spacing met with is 127 cm.$^{-1}$ for iodine. A band structure in the absorption spectrum of mercury with a spacing of 18 cm.$^{-1}$ was regarded as vibrational† but even this was unprecedented. The individual maxima in that case fulfilled the requirement of having heads, with definite degradation in one direction (to the red). The present ones do not show this feature, though it may possibly be concealed by the smallness of the scale.

It is worth while to note, as suggested to me by Mr. Jevons, that a spaced vibrational *sequence* due to transition between (*e.g.*) corresponding members of two almost equally special sets of vibrational levels, will itself have a spacing of smaller order of magnitude than the vibrational levels themselves.‡ But here again we encounter the difficulty that such sequences should be repeated at intervals along the spectrum, while the present structure is not.

Upon the whole, considering the unusual features of this band, and particularly its association with the apparent structureless band with a head at 2470, it is best to admit complete agnosticism as to its true nature.

Summary.

The following are some of the chief observations out of a rather miscellaneous collection.

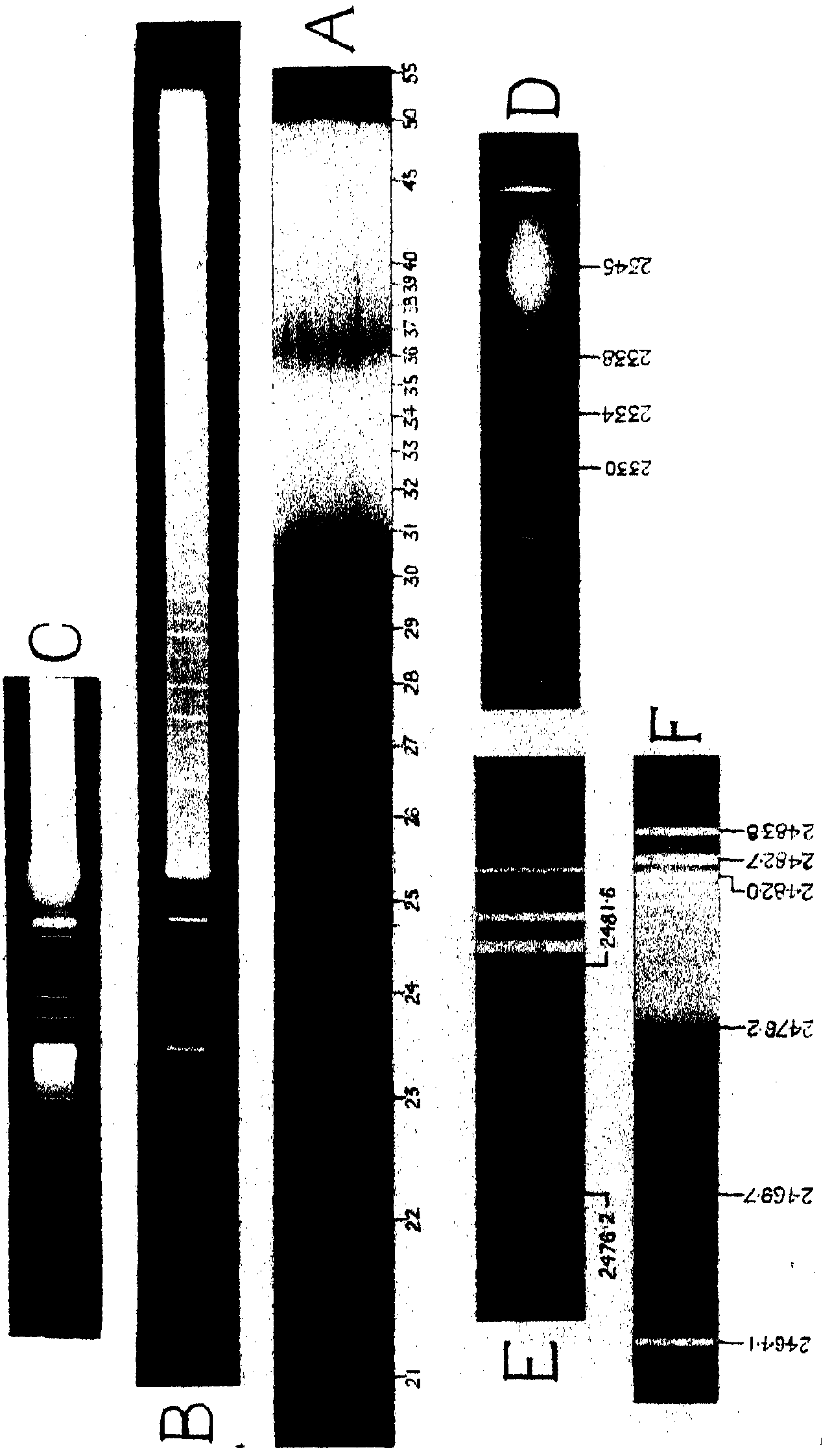
A mercury band spectrum is described which is excited by fluorescence, with the continuous hydrogen spectrum as a source. This spectrum is remarkable as showing in emission the band 2540 near the resonance line 2537, without the resonance line itself. It is observed also that the continuous spectrum on

---

* National Research Council U.S.A. Report on Molecular Spectra, December, 1926, p. 230.

† 'Roy. Soc. Proc.,' A, vol. 116, pp. 715, 717.

‡ Such sequences are well seen in the negative band spectrum of nitrogen. Reference may be made to a spectrum of the aurora which I published some time ago, 'Roy. Soc. Proc.,' A, vol. 101, p. 124 (Plate 2) (1922).


A
B
C
D
E
F
55
50
45
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
2345
2338
2334
2330
2481·6
2478·2
2483·8
2482·7
2482·0
2478·2
2469·7
2464·1

the short wave side of the bands 2345, 2338, etc., extends as far as 2150, and thus much beyond the position of the forbidden line 1 $^1S_0$—1 $^3P_2$ 2270 which was found under other conditions to limit this continuous spectrum.

The series of bands 2345, 2338, etc., is examined in emission with large resolving power. It is found to be free from the complications of a finer underlying structure which appear in absorption.

A remarkable structure is found in the band 2482 to 2476, described by earlier writers as continuous. The proper description of this structure in terms of the quantum theory is, however, enigmatic. This band does not appear in absorption. The various spectra are reproduced.

DESCRIPTION OF PLATE 4.

A.—Fluorescence spectrum of mercury vapour under excitation by the continuous spectrum of hydrogen. Note the band 2540 and the long extent of the stretch of continuous spectrum from 2345 towards short waves.

B.—Spectrum of mercury under electric discharge in fairly dense vapour, using external electrodes. Note the development of several bands of the series 2345, 2338, etc., of which only first appears in A.

C.—Same as B, but more lightly printed. Note the stretch of continuous spectrum from 2345 towards short waves, as in A.

D.—The band 2345, 2338 photographed in emission with large spectrograph. Iron comparison below.

E.—The bands 2482 to 2476 with the weaker band 2476 to 2470.

F.—The stronger band 2482 to 2476, showing fine structure.

The Study of the Specific Resistance of Bismuth Crystals and its Change in Strong Magnetic Fields and some Allied Problems.

By P. KAPITZA.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 30, 1928.)

[PLATES 5–8.]

Introduction.

It is well known that in a magnetic field bismuth shows a greater change of resistance than any other substance, and it is also known that in the case of a crystal this phenomenon varies very much with the orientation of the crystal. A great deal of literature exists on this subject.*

The general view of the phenomenon is that the increase of resistance is largest when the cleavage plane of the crystal is parallel to the magnetic field, and when the current is flowing perpendicular to it. It is also known that the resistance in a magnetic field increases very rapidly with decreasing temperature. A complication in all these phenomena arises through certain time lags. When a current is passed through bismuth placed in a magnetic field, the resistance at the first moment is large, and then gradually decreases to its final value. This time lag accounts for the fact, first discovered by Lenard, that bismuth has a larger resistance for alternating currents than for direct currents. This phenomenon also depends on the crystal state of the bismuth.

In the present experiments both these phenomena in bismuth crystals have been studied, using strong magnetic fields up to 300 kilogauss, obtained by the method recently described.† The magnetic field obtained by this method lasts for only about 1/100 second, so that a special method had to be developed for measuring the changes of resistance occurring in such a short time. As the space available in the coil for the experiment is only 2 c.c., very small samples of bismuth, in the form of rods, must be used. These rods cannot be successfully cut out from a large block of crystal, as it is well known that the surface is disturbed to a considerable depth by such operations. It was therefore decided to develop a method of growing bismuth crystals in the form of small rods, in which the crystal planes have a given orientation. This proved to be a more difficult task than had been anticipated. Several ways of attack were tried before a satisfactory method was found.

---

* Graetz, 'Handbook of Electricity,' vol. 3, p. 698, and vol. 4, p. 1026. See also Baedeker, 'Die electrischen erscheinungen in metallischen Leitern,' p. 94.

† Kapitza, 'Roy. Soc. Proc.,' A, vol. 115, p. 658 (1927).

In working out this method of growing rods of bismuth crystal, the author noticed certain phenomena which bear a close relation to those observed in the magnetic field and these will be discussed in some detail. This paper is accordingly divided into three parts, the first dealing with the method of growing bismuth crystals, the second with the methods of measurement, and the third with the magnetic phenomena.

Part I.—*The Growth of Crystal Rods with a Definite Orientation of the Crystal Planes and the Specific Resistance of Bismuth Crystals.*

(1) *Experiments in Growing Bismuth Crystals.*

Bismuth was one of the first metals to be obtained in large crystals, and it is considered to be one of the substances which crystallises most easily. Bismuth crystallises in the hexagonal system, and its crystallising symmetry is that of the rhombohedron, which approaches very closely to a cube. It has a trigonal axis, and perpendicular to it a perfect cleavage plane which is one of the four pseudo-octahedral planes. The other three pseudo-octahedral planes are also cleavage planes, but are not nearly so good as the one perpendicular to the axis. These cleavage planes are easily distinguished from each other. In a good crystal the perfect one has three sets of lines of equal strength intersecting at an angle of 60°. These lines run parallel to the lines of intersection of the three imperfect cleavage planes of the pseudo-octahedron. An imperfect cleavage plane has three sets of lines also, but one set (which is parallel to the perfect cleavage plane) is very strong, and the two sets parallel to the remaining imperfect cleavage planes are weak. The angle between the weak lines is 65° 28′, and between the strong and the weak 57° 16′.

Along the trigonal axis and perpendicular to it practically all the physical properties of bismuth are found to differ. Thus the electrical conductivity along the trigonal axis is smaller than that perpendicular to it. The same is true of the thermal conductivity, while the magnetic properties are also different.

For our experiments we required two sets of crystal rods—one with the perfect cleavage plane perpendicular to the length of the rod, *i.e.*, the axis of the crystal parallel to the axis of the rod; and the second with the crystallographic axis perpendicular to the length of the rod. For the various intermediate positions practically all the phenomena can be generally deduced by the well-known methods of tensor analysis.

The size of rod required for our experiments was about 3 to 5 mm. in length, with a cross section of about 1 square millimetre.

We first tried to obtain such rods by growing them in small glass tubes, of a diameter of 1 mm. filled with fused bismuth, and then slowly cooling them from one end—a method similar to that used by Obreimow and Schubnikow* and Bridgman.† The glass is then dissolved in hydrofluoric acid. In each of these rods the angle between the perfect cleavage plane and the axis of the rod varied from 0° to 60°, but not a single rod had the cleavage plane perpendicular to the axis of the rod. It is worth mentioning here that Bridgman‡ in all his numerous experiments on the growth of bismuth rods never obtained a larger angle than 45°.

We then tried a modification of Czochralski's§ method of growing crystals in which the rod is drawn from a crucible filled with molten bismuth. The rod is cooled and the bismuth is solidified slightly above the level of the molten bismuth in the crucible. The rods obtained in this way are fairly circular but of variable diameter. In order to obtain rods of a given diameter we modified this method, and instead of keeping the molten bismuth in a crucible we used a vertical pyrex tube of 5 mm. cross section and about 5 cm. in length. At the top end the tube was drawn to a nozzle of a diameter slightly larger than the required rod; at the bottom the tube was fitted with a glass piston with asbestos filling. The tube was wound with a few layers of nichrome wire, with asbestos insulation, and a small current through this winding was sufficient to keep the bismuth in a molten state. The crystal was grown from the nozzle by bringing a small piece of bismuth into contact with the meniscus of molten bismuth on the nozzle, and then very uniformly and slowly pulling it vertically up. By means of a simple mechanical device it was arranged that the asbestos piston was moving quite uniformly with the growing crystal rods, but with a reduced velocity, and thus forcing the necessary amount of bismuth through the nozzle. To obtain uniformity of motion, an electric motor with a reduction gear was used.

In this way it was possible to obtain uniform rods of 1 sq. mm. section, but not a single one of these rods had the angle of the perfect cleavage plane closer than 30° to the perpendicular plane of the rod.

---

* 'Z. f. Physik,' vol. 25, p. 31 (1924).

† 'Proc. Amer. Acad. Arts Sci.,' vol. 60, p. 307 (1925).

‡ *Loc. cit.*, p. 349.

§ 'Z. f. Phys. Chem.,' vol. 92, p. 219 (1917).

Georgieff and Schmid* used the Czochralski method for growing rods of bismuth crystal, but in all their specimens they did not obtain a closer approach of the cleavage plane to the perpendicular than I obtained in these experiments.

Further attempts were made to force the rods to grow with the cleavage plane perpendicular to the length by growing them from a definite plane. This was done in the following way. On the cleavage plane of a large crystal, grown in a crucible and broken along the cleavage plane, was placed vertically a small 1 mm. bore glass tube filled with bismuth. The glass tube was surrounded by a short platinum spiral of a slightly larger diameter, which could be heated by a current, and by means of clockwork, moved, without touching the tube, along its length. The spiral was first placed on the bottom of the rod, and the current was passed until the bismuth melted, and the rod sank slightly into the big crystal. The bismuth could not flow out of the glass tube as it was held back by the barometric pressure. Clockwork was then set in motion and the spiral was gradually lifted, melting the bismuth higher up in the tube and letting it solidify at the bottom. In this way it was expected that a large crystal would grow in the glass tube with the required orientation.

The other method of drawing the crystal rod from the nozzle was used for the same experiment. In this case the small bismuth rod from which the crystal starts to grow was cut out of a big piece of crystal with its cleavage plane perpendicular to the length.

In neither case was a positive result obtained. What actually happened was that in the rod grown in this way a cleavage plane was perpendicular to the rod; this was not the perfect cleavage plane, but one of the remaining three planes of the pseudo-octahedron. This was easily established by looking at the plane under the microscope and also from the specific resistance measured along the rod.

From this experiment it is evident that there was a quite definite cause preventing the crystal from growing along its trigonal axis. Later we shall give what is probably the true explanation of these interesting results, but at first we attempted to explain them in the following way.

It is known that bismuth crystals should grow best in a direction along the perfect cleavage plane, as in this direction the crystal has the maximum heat conductivity and a closer packing of the atoms. As in the two methods of growing the crystals the thermal gradient was along the rod, and as the rod is growing along the gradient, the above-mentioned properties of the bismuth crystal could conceivably account for the failure to grow rods with the perfect cleavage plane perpendicular to the length. An attempt was therefore made to set up a temperature gradient more or less perpendicular to the axis of the rod. This was done in our first method by placing the platinum spiral on one side of the glass rod and a small copper plate on the other side. In the second method the perpendicular gradient was also established by means of a platinum spiral and a copper plate placed on opposite sides of the nozzle. In both cases, however, after making a great number of attempts, we were unable to obtain bismuth rods having a perfect cleavage plane perpendicular to the axis of the rod.

---

* 'Z. f. Physik,' vol. 36, p. 759 (1926).

The results were exactly similar to the previous ones; if a rod was started from a perfect cleavage plane of a given crystal, the perfect cleavage plane would change in the grown rod to an imperfect one, and the angle between the perpendicular plane of the rod and the perfect cleavage plane would never be less than 30°. It was therefore clear that another factor was preventing the growth of the crystal with the perfect cleavage plane perpendicular to the axis, and the only possible assumption was that this factor was the strain set up in the crystal during its solidification. In the first method with the glass tube, the strain may be due to the fact that during its solidification bismuth increases in volume by about 3 per cent. This can easily be noted by the fact that in cases where the glass tube had thin walls it cracked during the process of the solidification of bismuth. In the second method, as the bismuth solidified only 3 to 4 mm. from the edges of the nozzle, the strain was longitudinal, due to the weight of the liquid part of the rod.

It was therefore decided to use a method by which no strain is set up during the process of the growth of the crystal. This was accomplished in the following manner. A copper plate *a* (fig. 1), about 25 cm. long, 1 cm. thick and 5 cm. broad, had one end fastened to a stand *b*, and on the other end was wound, for about 5 cm., a heater *c* of nichrome wire insulated with asbestos. When the current was sent through the wire the plate was fairly uniformly heated, having a small gradient of temperature from the heated end towards the supported end. A bismuth rod *d*, with a diameter of 1 to 2 mm., was placed on a glass or quartz plate in the middle part of the copper plate *a*. In order to prevent the cooling of the rod by accidental air currents, it was protected by glass plates which were placed at the sides and on the top of the rod, as is shown in cross section on the small drawing (A). A steady current of 2·5 amperes at 60 volts, from accumulators, was sent through the winding of the heater and the plate was heated to a temperature at which the bismuth began to melt, this requiring about 30 to 45 minutes. When liquid, as a result of surface tension and a small layer of oxide, the rod kept its circular shape very well and was scarcely deformed at all. When the rod was melted the cooler end was pulled out by touching

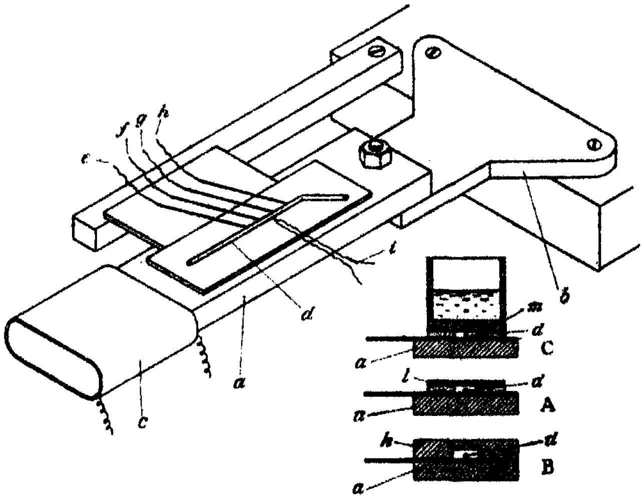


FIG. 1.

it with a small glass rod so as to make a small sharp appendix. (This appendix is necessary for the same reason as in Obreimow's method of growing crystals in glass tubes, where the end of the glass tube was drawn to a capillary, this making it much more certain that the crystal will commence growing in one spot and will produce a monocrystalline rod.) After this the current was very slowly diminished and the rod commenced to solidify from the sharp end of the appendix, crystal rods being obtained in this way. After a few experiments it was found that in such a way it was possible to obtain rods with the cleavage plane orientated at any angle relative to the axis of the rod. In order to make certain that the rod is monocrystalline the solidification time has to be not less than half an hour for a rod 7 to 10 cm. long.

To obtain the desired orientation of the cleavage plane in the crystal, the following method was adopted. A piece of rod, about 2 or 3 cm. in length (obtained by the method described above), was broken off from a long rod along the cleavage plane. A second bismuth rod was then melted on to the first rod in such a way that the cleavage plane formed the desired angle with the axis of the new rod. This was done in the following way. The first rod was fixed on to a glass plate by means of clamps and the second rod was then approached at the required angle, which was drawn on a paper placed under the glass plate. Then a small gas flame, about 3 mm. long, obtained from the end of a capillary glass tube, was brought to the point of contact between these two rods. The bismuth commenced to melt and the rods were slightly pressed to each other. This melting process was slightly tricky, as sometimes a layer of oxide may come in between the joint and spoil the experiment. In order to avoid this it is necessary to melt the rods with the middle part of the flame only, where there is an excess of gas, as this prevents oxidisation. The two rods, melted together in this way, were placed as before, on the glass plate on top of the copper plate, in such a manner that the already crystalline part of the rod was on the cooled end of the plate *a* as shown on fig. 1. The rods were protected from any air currents, except that the outside half of the already crystalline rod was uncovered. This increased the gradient of temperature in this part of the rod when the plate was heated.

The process of growing the crystal is very similar to the previous one. The copper plate is heated up and the bismuth is melted, but in this case it is done in such a way that only half of the already crystalline rod is melted.

It is worthy of mention that the heating of the rod must be done very carefully, as the gradient of temperature is very small, and it is very easy to melt the whole of the rod, including the part from which the crystal is growing. This difficulty can be avoided by fixing a thermo-couple near the rod and recording the temperature accurately, or by placing small pieces of bismuth on various parts of the copper plate near the rod, and touching them with glass rods to see when they have melted. Otherwise it is very difficult to see when the rod begins to liquefy. When the rod was slowly cooled it grew in a crystal in which the cleavage plane was orientated towards the axis of the rod at the necessary angle. In this way, crystals of any given orientation of the cleavage plane towards the axis of the rod could be obtained without difficulty, with an accuracy of 1° to 2°.

The bismuth used for our experiments must be of high chemical purity. We used bismuth obtained from Kahlbaum and from Hartmann and Braun. The latter firm supplied bismuth already drawn in wires of the required diameter of 1 mm. From Kahlbaum's bismuth, supplied in lump form, rods of the necessary diameter can easily be made. This can be done by pressing bismuth, in a warm condition, through a hole, or by filling a small thin walled glass tube with molten bismuth and dissolving the glass in hydrofluoric acid after solidification. From experiments in magnetic fields it was seen that Hartmann and Braun's bismuth has a higher degree of purity than that of Kahlbaum and is therefore more suitable for our experiments.

(2) *Specific Resistance of Bismuth Crystals perpendicular to the Cleavage Plane.*

It is well known that the specific resistance of bismuth crystals is greater perpendicular to the perfect cleavage plane than parallel to it, but the actual values of the specific resistance in either direction obtained by various authors differ considerably. This is particularly noticeable in the data for the specific resistance perpendicular to the cleavage plane. We find that the values range between $1 \cdot 6$ and $2 \cdot 06 \,.\, 10^{-4}$ (Matteucci, Everdingen, Voigt, Lownds, Borelius and Lindh). In this paragraph we shall describe the results of our measurement of the specific resistance in this direction, and we propose to give an explanation of the discrepancies observed in previous investigations.

The measurement of the specific resistance of crystal rods, grown by our method with a definite orientation of the axis, was carried out by measuring the resistance of a rod 2 or 3 cm. long by a potentiometer method. Two readings were taken with the current in opposite directions through the rod so as to exclude the thermo-electric effect, and the mean value taken. The cross section was determined from the weight and the length of the rod by taking the specific gravity of bismuth to be 9·80.* In our measurements we did not aim at an accuracy greater than 1 per cent.

To our surprise we found that the specific resistance perpendicular to the cleavage plane varied with different specimens of crystal rods. The lowest obtained was $1 \cdot 39 \,.\, 10^{-4}$ and the highest $1 \cdot 6 \,.\, 10^{-4}$. The fact that crystal rods grown in approximately the same way, and from the same material, should show such a difference in specific resistance was remarkable.

We found some hints as to the explanation of this phenomenon in the work of Borelius and Lindh.† These authors observed that a specimen of bismuth crystal, cut out from a large crystal, perpendicular to the cleavage plane, diminished its resistance when placed in the holder of their apparatus, where the compression was mainly axial.

We therefore decided to take different samples of our crystals and study the change of the specific resistance when compressed. The arrangement used is shown on fig. 2.

The bismuth rod *a*, which is grown with its cleavage plane perpendicular to its length, is placed vertically and rests on a support *b* which is fixed to a vertical ebonite plate *c*. This plate is fastened to the table. The bottom end of the bismuth rod *a* rests on a platform *d*. This platform, when lifted, will compress the crystal. To ensure a parallel motion of the platform it has a guide, made of a small rod, which slides through a cylindrical hole in the brass piece *e*, fixed to the plate *c*. The compression is transmitted to the crystal by means of an arrangement shown on the same figure, and which consists of a lever *f*, two strings *g* and a small horizontal plate *h*. When the lever is loaded the crystal is compressed. This little apparatus has to be rather carefully made, and the rod accurately adjusted, as it is very easy to bend the crystal during the compression and then the experiment is spoilt. It is also necessary for the diameter of the bismuth specimen to be correctly proportioned to its length. We find that ratio of the diameter to the length should be about 1 : 6, to prevent it from bending.

---

* We took this value from our tables and also verified it by direct measurement.

† 'Ann. der Physik,' vol. 51, p. 613 (1916).

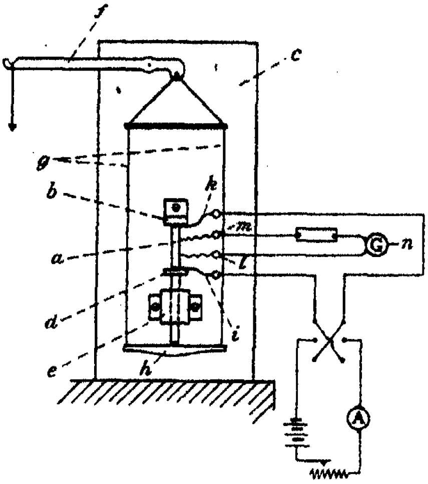


FIG. 2.

The resistance was obtained from the potential drop in the rod when a current was passed. The current leads consisted of two fine copper strips, *i* and *k*, soldered by means of a soft solder (Wood's metal or Newton's metal) to the ends of the bismuth rod *a*. The two potential leads, *l* and *m*, are of fine silver wire ($d = 0 \cdot 1$ mm.), and are welded to the rod by means of a spark from a condenser. This is done by a well-known method. The bismuth rod is connected to one end of a condenser (50 microfarads) through a small resistance, and the silver wire to the other end. The condenser is charged to 20 to 30 volts, and it is then sufficient to touch the bismuth rod at the place where we wish to make the joint with the end of the wire. A small spark occurs and the wire is well joined. A good point joint is obtained in this way without spoiling the crystal. The potential drop was measured by a high resistance (about 500 ohms) galvanometer *n*. To exclude the influence of the thermal e.m.fs. two readings are necessary with opposite directions of the power current.

On fig. 3 the experimental results are shown for several rods which have different initial resistances. The abscissæ represent the load in grams per square millimetre, and the ordinates the specific resistance. For clearness the three curves are shown one above the other. The bottom curve (i) relates to a compression of a rod made from Hartmann and Braun bismuth and grown in the way described in the previous paragraph. The rod of curve (ii) is made in a similar way, but of bismuth from Kahlbaum. Curve (iii) refers to a rod made from Kahlbaum bismuth, but cut out of a large crystal grown in a crucible.

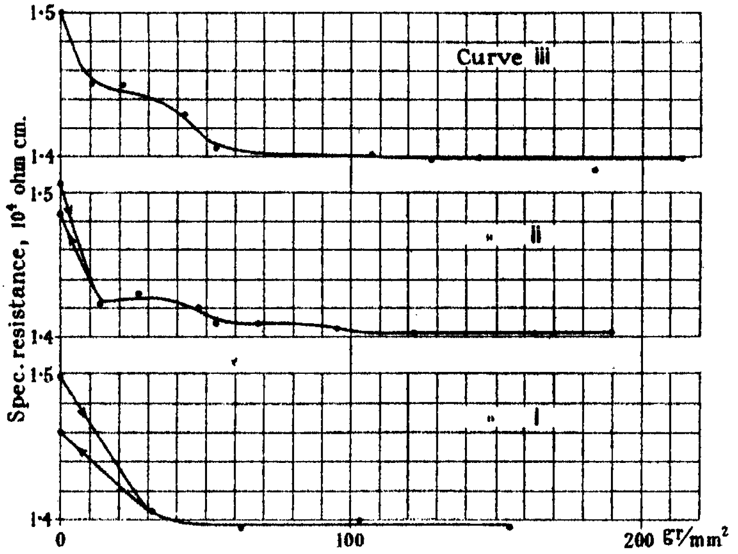


FIG. 3.

It is seen that the resistance of various rods in a compressed state approaches $1 \cdot 39$ to $1 \cdot 4 . 10^{-4}$, within the limits of experimental error. After the compression reached the value of about 50 grams per square millimetre the resistance of the rods did not change appreciably.

These experiments suggest that bismuth rods have " cracks " of appreciable resistance along their cleavage plane, and that by means of the compression these cracks close and thus diminish the resistance. The limiting value of the specific resistance, $1 \cdot 39 . 10^{-4}$, is reached when all the cracks have closed. This view is strongly supported by the general character of the change of the resistance of bismuth as it approaches the limiting value under compression. The diminution of the resistance for different specimens occurs in various ways and sometimes occurs in jumps, as is seen on fig. 3. When the load is taken off the resistance increases again, but there is generally a tendency for the crystal to retain the smaller resistance, and the final resistance of bismuth, after unloading, is smaller than the initial one (see curves 1 and 2). A slight pull on the crystal brings the resistance to the initial value, and sometimes makes it higher.

From this experiment we came to the conclusion that the true specific resistance of the bismuth crystal perpendicular to the cleavage plane is $1 \cdot 39 . 10^{-4}$, and the higher value observed in different specimens, both by myself and other authors, is due to the cracks which develop in bismuth, not only during the process of cutting out the rod, but during the growth of the crystal. This development of cracks during the formation of a crystal is a very interesting phenomenon indeed, and we shall discuss the question later in more detail.

At this stage, mention should be made of the recent work of Bridgman,* involving a very careful determination of the specific resistance of a large number of monocrystalline bismuth rods grown in glass tubes. As already mentioned, by this method of growing crystals, Bridgman could only obtain rods having angles ranging from 90° to 45° between the trigonal axis and the axis of the rod. He found that the specific resistance varied for the same orientation of the cleavage plane in the rod. He chose three rods having the smallest resistance, with the perfect cleavage planes inclined to the axis of the rod at angles of 43°, 27° and 4°, and by extrapolation found the specific resistance for a rod in which the cleavage plane is perpendicular to the axis of the rod, obtaining the value of $1 \cdot 38 . 10^{-4}$ which, taking into account the manner in which it was obtained, agrees very well with that found in our experiments.

(3) *The Specific Resistance of Bismuth Crystals parallel to the Perfect Cleavage Plane.*

The results obtained by different authors for the specific resistance of bismuth crystals parallel to the perfect cleavage plane vary to a much less extent than in the case of the specific resistance perpendicular to the cleavage plane. At room temperature the values given range between $1 \cdot 09 . 10^{-4}$ and $1 \cdot 14 . 10^{-4}$.

---

* 'Proc. Amer. Acad. Arts Sci.,' vol. 60, p. 360 (1925).

For our crystals the measurements for a crystal grown on the copper plate (shown in the apparatus given in fig. 1) gave values ranging between 1·07 and $1{\cdot}14 . 10^{-4}$, this being even a larger variation than that obtained in the results of previous investigators. In this case it is hardly possible to attribute the variation to cracks. No compression experiments are possible with bismuth rods grown in this direction, as they are too soft and bend very easily. The reason for this variation of the specific resistance is probably to be found, as will be seen from later experiments, in the fact that the trigonal axis changes its direction in certain very small parts of the bismuth crystal. This explanation leads us to believe that the lower value is the correct one, and this also does not differ greatly from the value given by Bridgman* ($1{\cdot}09 . 10^{-4}$).

(4) *The Influence of Impurities and of the Temperature Gradient on the Formation of Bismuth Crystals.*

The variation in the specific resistance in different directions of the bismuth crystal, discussed in sections 2 and 3, is ascribed to the imperfection of the crystal, which is due to cracks formed along the perfect cleavage plane or to local variations of the crystallographic axis. Since in our method of growing the crystal all outside strains are eliminated and the specific resistance is measured without any strains being produced by the process of cutting it from a large lump of crystal, or removing it from a glass tube, as in previous investigations, we must conclude that this imperfection occurs during the process of the formation of the crystal.

In this case the imperfection may be due to two sources; first, the way in which the crystal is cooled, and secondly, the chemical purity of the bismuth. The study of these two factors will be described at length in this section, and it will be seen that in this way some light is thrown on the problem of the source of the cracks. In the first place the effect of the influence of the temperature gradient on the growth of the crystal is discussed.

In order to vary the temperature gradient perpendicular to the axis of the bismuth rod, we used the arrangements shown on fig. 1 by the small drawings A, B and C. These drawings represent different arrangements used for covering the bismuth rod *d* which is resting on the copper plate *a*. In the arrangement shown on fig. 1 B, the rod is covered by means of a solid copper plate with a channel. It is clear that here the temperature gradient perpendicular to the rod is very small. If a glass plate *l*, fig. 1 A, is used, instead of a copper plate, the temperature gradient is increased, since the rod is cooled more rapidly from the top. Finally, if in the place of a glass plate we have a small copper vessel containing water *m*, fig. 1 C, the temperature gradient is still more increased.

In order to alter the gradient along the crystal rod we varied the cross section of the copper plate *a*. It was found, however, that such a method did not give a sufficient increase of the temperature gradient along the rod, and therefore an arrangement, shown on fig. 4, was used, whereby it was possible to obtain a very large temperature gradient along the rod. This apparatus consisted of a solid copper plate *a* having a square opening in the middle. Above this opening was placed a small spiral *c* made of platinum wire. Through this spiral a pyrex or quartz tube *f* can be moved by means of a frame *e* which is sliding on the copper plate *a*. The tube *f* is resting on the copper plate *a*, and is directed in its motion by passing through a hole made in a small copper appendix *d* which is fixed to the copper plate *a*. The bismuth rod *b* is resting quite freely in the tube *f*. The crystal is grown in this apparatus in the following way.

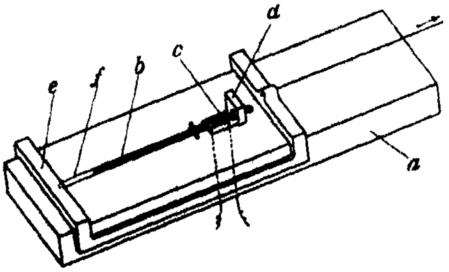


FIG. 4.

A current is passed through the spiral, which heats up the bismuth rod locally to a temperature which is sufficient to melt the bismuth. By means of clockwork the tube is pulled very slowly and uniformly through the platinum spiral and fresh portions of the bismuth rod enter the spiral and are melted. On the other hand the bismuth which is already melted is cooled and it then enters the hole in the copper appendix *d*. In this way the whole rod is gradually turned into a crystal. It is obvious that by adjusting the distance between the copper appendix *d* and the spiral *c* any very large temperature gradient may be obtained along the rod. By means of an arrangement (not shown in fig. 4) the spiral and the heated part of the tube were protected from air currents.

To obtain rods with a given orientation of the cleavage plane from crystals grown with this apparatus, the following method was adopted. A short rod with the required orientation of the cleavage plane was prepared on the copper plate of the old apparatus shown on fig. 1, in the way already described in section 2. This rod was then placed in the tube in such a way that the join was slightly to the left of the spiral, and the already crystalline portion passed through the hole in the copper appendix *d* in the cool region. It is therefore evident that in this way all the rod will be turned into a crystal having the same orientation of the crystal axis as in the appendix.

Finally, in order to have a very small temperature gradient in all directions, the simple arrangement in fig. 5 was used. This consists of a solid copper block, in the form of a cylinder *a*, with a small central bore *f*. This cylinder was warmed by means of an electric heater *b*. In order to have the cylinder at a very even temperature it was wrapped up in asbestos. The bismuth rod *d* was freely placed in the glass tube *c* and inserted in the hole *f* of the cylinder, as shown on fig. 5. In all this apparatus it was found that the current had to be supplied by accumulators in order to obtain the requisite steadiness.

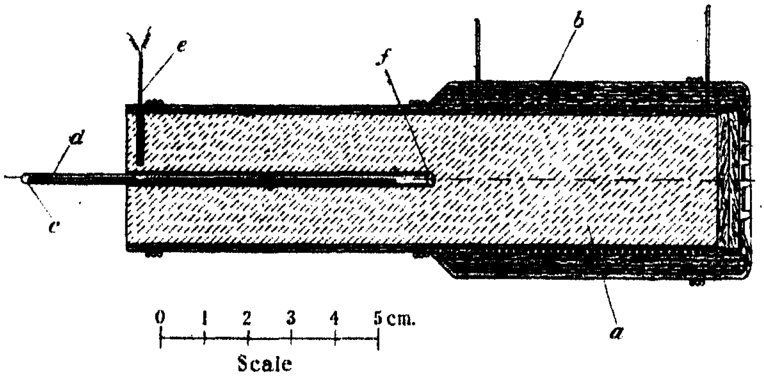


Fig. 5.

To obtain the required orientation of the cleavage plane, the bismuth rod was prepared in exactly the same way as in the previous case. The method of growing the crystal was very simple. The cylinder *a* was heated up to a temperature such that the bismuth rod was melted a little farther from the place where it was joined to the crystalline rod and was then very slowly cooled down by reducing the current of the heater. The moment at which the temperature was sufficient to melt the bismuth was recorded by the thermocouple *e*, which was placed just above the point where the two rods joined. It is worth mentioning that in this method, as in the previous ones, for obtaining good crystals, it is advisable to make one or two reductions in section on the rod (*see* fig. 5). The rod is then certain to be monocrystalline.

In growing the various bismuth crystals by these three methods, and in varying the orientation of the perfect cleavage plane relative to the temperature gradient, the following phenomena were observed.

It is known that* the flexibility of monocrystalline bismuth rods varies with the orientation of the perfect cleavage plane relative to the axis of the rod. When the perfect cleavage plane is parallel to the axis of the rod, the rod is fairly flexible, but the flexibility diminishes very rapidly as the cleavage plane approaches a perpendicular orientation to the axis of the rod. Rods in which the cleavage plane is perpendicular to the axis are stated to be very brittle, but in our methods of growing crystals we observed that the flexibility of rods with the same inclination of the cleavage plane, grown with a different temperature gradient, varies greatly. In some instances it was possible to obtain very flexible bismuth rods in which the perfect cleavage plane was only a few degrees from being perpendicular to the axis of the rods. For example, one rod with its perfect cleavage plane only 15° out of the perpendicular position could be bent in a radius of curvature of 3 mm. as easily as a rod made of tin. Another, with the same orientation of the cleavage plane, was so brittle that, when falling on to a table from a height of a few centimetres, it broke into many small pieces along its cleavage plane. The difference between the flexible and the brittle bismuth can also be seen from the following facts. During bending, the less flexible rods, if held close to the ear, emitted cracking sounds, whereas others, being more flexible, did not. It was easily established that the more flexible the rod the less cracking sounds were produced on bending.

It was observed that there is a close relation between the flexibility of the rod and its specific resistance. All the flexible rods gave the same specific resistance, and this resistance corresponded to the two lowest values given for the crystal. It was a curious fact that even after a flexible rod had been considerably bent if no cracking sound was heard the specific resistance increased but slightly. If, however, cracking sounds were heard the resistance increased very rapidly. This suggests that these sounds are probably caused by the development of the cracks and imperfections already existing in the lattice of the bismuth crystals. As regards cleaving the rod with a knife, the softer specimens at room temperature would not cleave at all. The knife would cut almost through it before it broke. The brittle specimens were somewhat better, but frequently a small piece of the rod would be cut through without showing a cleavage plane at all.

We shall discuss later the origin of this difference in flexibility of bismuth rods, but will first describe the effect of the temperature gradient and impurities on the flexibility of bismuth rods. This was at first difficult to trace, but after a large number of experiments we came to the following conclusion.

---

* See Georgieff & Schmidt, *loc. cit.*

To obtain the most flexible bismuth rod during the formation of the crystal, a very sharp temperature gradient perpendicular to the perfect cleavage plane is required. We will first consider the case in which the growing crystal has its cleavage plane perpendicular to the axis of the rod. If we grow such crystals in the apparatus shown on fig. 4, with the platinum spiral, we obtain fairly flexible crystals, and it is evident that in this case the gradient will be along the rod and perpendicular to the cleavage plane. The specific resistance of these crystals was always the same and was equal to $1 \cdot 38$ to $1 \cdot 39 \,.\, 10^{-4}$. This resistance corresponded exactly to the one observed in the rods when they were compressed in the way described in section 2. This means that the rod has no cracks. The very flexible crystals which, as mentioned previously, could be bent to a radius of 3 mm., were obtained by this method, but only occasionally. In none of these flexible rods was the cleavage plane exactly perpendicular to the axis of the rod, but in the case just mentioned we obtained a crystal in which it was only 15° from perpendicularity. The most probable reason for this is that in order to obtain the most flexible crystals it is necessary to have the temperature gradient very closely perpendicular to the perfect cleavage plane. In the apparatus shown on fig. 4 this condition was only approximately fulfilled, as the cooling round the rod was not quite symmetrical, and it was only in particularly "lucky" experiments that the condition of perpendicularity was fulfilled.

When crystals were grown with this orientation on a copper plate (fig. 1) with a transverse temperature gradient, they were always brittle and their specific resistance varied and was always higher than $1 \cdot 43 \,.\, 10^{-4}$. When compressed, this resistance was reduced to the normal value of $1 \cdot 38 \,.\, 10^{-4}$, as was mentioned in section 2.

In growing crystals under the most uniform temperature conditions, as produced with the copper cylinder, in the apparatus shown on fig. 5, we obtained the most brittle variety with its cleavage plane perpendicular to its length. Some specimens were accidentally dropped on to the table from a height of a few centimetres and they broke into many small pieces, which were only 2 or 3 mm. long. These small pieces were not very brittle and it was even possible to bend some of them slightly. The specific resistance of these crystals, when carefully handled, would give the normal value of $1 \cdot 38 \,.\, 10^{-4}$, but a slight strain would at once increase this resistance.

If the crystal was broken by bending, it was observed that the cleavage plane in the hard specimens was very perfect and gave a perfect mirror-like reflection. In flexible specimens the cleavage plane was slightly uneven, but this unevenness of the surface was not due to the crystal itself, but was produced during the breaking up, as it depended very much on the way in which the crystal was handled. For instance, if the crystal was cooled by liquid air, before being broken, the cleaved plane was quite good. (The X-ray analysis described later in the paper also confirms the statement that a change of the cleavage plane is produced by bending the crystal, as the lattice of the flexible crystal shows no traces of being bent.)

When growing crystals with the cleavage plane parallel to the axis of the rod, we were able to obtain soft flexible rods only when they were grown on a plate (fig. 1) with a rather strong vertical gradient, and when the cleavage plane was almost parallel to the copper plate. The flexible crystals had the normal resistance of 1·07 to 1·08, and in some cases these were so flexible that if one end of the rod was raised it began to bend under its own weight. They could be bent considerably without the least cracking noise being heard. It was found, however, that it was impossible to obtain an entirely flexible rod with its cleavage plane parallel to the length, just as previously we could not obtain one with its plane perpendicular to the length. In the best flexible specimens the angle was 20° out of parallelism with the axis of the rod. The explanation is probably similar to that given for the previous case.

When in the same apparatus the crystal was grown with its cleavage plane perpendicular to the copper plate, the rod was never flexible and the specific resistance lay between 1·10 and $1{\cdot}14 \,.\, 10^{-4}$. Similarly, in growing this type of rod in the apparatus with the platinum spiral, shown on fig. 4, it was found that it was never flexible, and the resistance varied in different cases from 1·10 to $1{\cdot}12 \,.\, 10^{-4}$.

Finally, when the rod was grown in the apparatus, shown in fig. 5, with the most uniform temperature, it always had the smallest resistance, 1·07 to $1{\cdot}08 \,.\, 10^{-4}$, but was invariably not flexible.

The speed of growth of the crystals in all these cases appears to affect the crystalline character of the rod very little, provided that it does not exceed a certain limiting value. It was also noticed that the speed with which these crystals are grown, with the perpendicular orientation of the perfect cleavage plane, must be less than for the rod with the perfect cleavage plane parallel to its axis. In the apparatus of fig. 4, for the cleavage plane with perpendicular orientation, this speed must be no more than 1 mm. per minute. For the other orientations of the cleavage plane, it may be 7 mm. per minute or even greater.

If the speed is higher, the perfect cleavage plane changes its orientation, or the rod is not monocrystalline. In growing crystals on the copper plate of the apparatus of fig. 1, the speed can be much higher, as in this case the crystal is now growing along the rod, but in a transverse direction. The surface on which the atoms arrange themselves in a lattice is larger and consequently they arrange themselves much more quickly.

The second question we studied was the influence of the impurities on the growth of the crystal. In the first instance we made certain that the variation in the flexibility of the crystal was not caused by the slight variation in the amount of impurity which probably existed in different specimens, by taking a bismuth crystal rod which was flexible, and regrowing half of it at a different temperature gradient. It was then possible to have the first half of the rod flexible and the newly grown part brittle.

In general, we found that more flexible rods were produced with greater ease from the less pure bismuth. We judge the purity of the bismuth from the specific resistance and from some data on the measurements of the change of resistance in a magnetic field, which will be described later.

We also experimented in order to determine the influence of the gases which are always found in the metals. For this purpose we took a rod of bismuth and placed it in an evacuated quartz tube which was made red hot. We kept the bismuth at this temperature for about 1 hour, and pumped the gas out of the tube during the whole of this time. During this process, precautions were taken to prevent the bismuth rod from being broken into drops, by placing round it properly shaped quartz rods. After cooling, the bismuth rod was placed in the apparatus, fig. 4, with the platinum spiral, modifications being made in the apparatus by which it was possible to keep the tube with the rod immobile, and move the spiral. This allowed us to keep the tube evacuated while the crystal was grown in vacuum. No appreciable difference in the flexibility and specific resistance was observed between rods grown in this way and those grown by the usual method. This suggests that the dissolved gases have no great influence on the development of the cracks. From all these experiments we conclude that the imperfections in the lattice are primarily due to the direction of the temperature gradient relative to the orientation of the crystal lattice in the place where the crystal is growing.

(5) *The Origin of the Cracks.*

It seemed important to trace the moment at which the cracks appear during the growth of the crystal, as this would indicate their origin. We expected to find the moment at which the cracks appeared indicated by a sudden change in the resistance of the bismuth during the process of growth. For this purpose we measured the resistance of the crystal during the process of its growth, using the arrangement shown on fig. 1. Four leads, *e*, *f*, *g* and *h*, were soldered to the bismuth rod. The two middle ones, *f* and *g*, were the potential leads, and were placed 10 mm. apart. The two end ones, *e* and *h*, were the power leads. The experiment consisted of melting and solidifying the rod by the method described in section 1, and growing the crystals with a definite orientation of the cleavage plane. The resistance of the bismuth rod during crystallisation was measured by the same method as that used in the compression experiment, described in section 2, only in this case large thermal e.m.fs. occurred and three successive readings were necessary for each given point with three successively changed directions of the current, the mean value being taken in the appropriate way. These experiments looked very simple, but for a long time they were quite unsuccessful, as it was very difficult to find material to form the leads to the rod. We tried to use silver and platinum wire, but these formed an alloy with the bismuth, and during the process of solidification they broke off. The problem was solved by making the leads of fine bismuth wire, manufactured by Hartmann and Braun, of 0·2 mm. in diameter, and welding them by a small flame to the rod. The wires were welded at an acute angle to the bismuth rod (as shown on fig. 1) in such a way that they could not solidify before the place at which they joined with the rod. In order to measure the temperature a constantan-iron thermo-couple *a*, fig. 1, made of fine wires, was placed between the potential leads at a distance of ½ mm. from the bismuth rod. The e.m.f. of the thermo-couple was measured by means of a potentiometer.

Curve i obtained in this experiment, and shown on fig. 6, relates to a bismuth rod which is solidified, and has its cleavage plane perpendicular to its length. The ordinates represent the resistance of the bismuth relative to the resistance in the melted state, and the abscissæ the temperature in arbitrary units. We see from this curve that bismuth at the point about 13·7, where it begins to solidify, increases its resistance practically at once by 2·64 times. The resistance then gradually commences to diminish and reaches at 16° a value of 1·05 times the resistance in the liquid state.

We took such curves, cooling the bismuth with various temperature gradients across the rods by the different arrangements, as already described and shown on fig. 1, by A, B and C, but in each case we obtained practically the same shaped smooth curve without any sharp bends. There was only a very slight difference in the shape of the maximum which was sometimes flat or sharp, but this was easily accounted for by the fact that the bismuth between the potential contacts does not solidify at once. The final resistance of bismuth crystals relative to that of the liquid state varied from 1·04 to 1·08, differing amongst themselves by about 4 per cent. In each experiment a piece of the rod was cut off from the end after the potential contact *g* and examined for flexibility and specific resistance.*

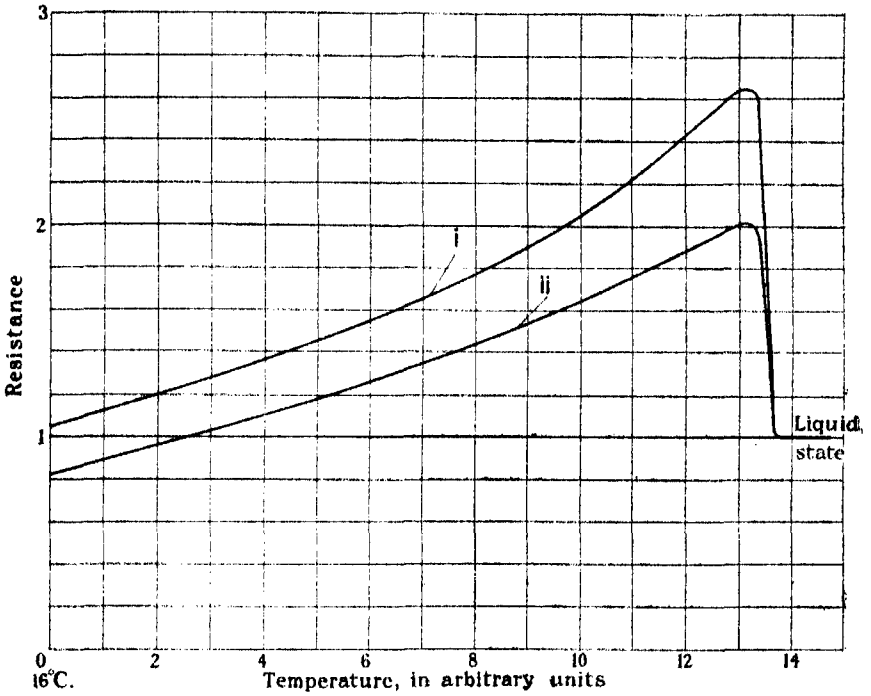


FIG. 6.

The crystal rod of curve i had a specific resistance of $1{\cdot}45 \,.\, 10^{-4}$ and the cleavage plane was inclined at 84° to the axis of the rod.

For other rods the specific resistance was found to vary between 1·43 and $1{\cdot}51 \,.\, 10^{-4}$, This variation is larger than was to be expected from the measurements during the solidification, where, as already stated, the variation was only 4 per cent. This is probably due to the fact that in handling the rod during measurement, some of the cracks widen.

---

* In order not to destroy the rather fragile contacts of the bismuth rod with the small wires, the cutting up was done by melting the bismuth rod at one point with a small flame.

In the case of the crystal grown with its cleavage plane parallel to the axis, we obtained curves which were also very smooth. One of these, curve ii, is shown in fig. 6. The maximum resistance in this case was 2·01 times that of the liquid state, and the final resistance at room temperature (16°) was 0·825. The specific resistance of the sample cut off was $1 \cdot 14 \, . \, 10^{-4}$ and the cleavage plane made an angle of 1° with the axis of the rod.

The smoothness of the curves and the absence of any sharp bends, together with the fact that the resistance increases with decreasing temperature, only at the moment when the crystal is solidified, indicates that the development of the cracks probably occurs almost at the moment when the bismuth passes from the liquid to the solid state.

It is interesting to note that if we take the specific resistance of liquid bismuth to be $1 \cdot 28 \, . \, 10^{-4}$, the resistance of the bismuth perpendicular to its cleavage plane would be, from these curves $1 \cdot 28 \times 1 \cdot 05 = 1 \cdot 34 \, . \, 10^{-4}$, and the specific resistance for rods with the cleavage plane parallel to the axis of the rod would be $1 \cdot 28 \, . \, 10^{-4} \, 0 \cdot 825 = 1 \cdot 055$. In both cases the values thus obtained are smaller than those afterwards measured in the samples cut off. This can be partly explained by the fact that the volume of the rod changes when the rod is cooled, but the difference seems to be rather too large to be explained entirely in this way. It is possibly due to an error in the determination of the specific resistance of liquid bismuth. This datum was taken from tables and represents the results of only one experiment made many years ago.

(6) *The Character of the Imperfection of the Bismuth Crystals.*

From the previous discussion it is seen that the bismuth crystals which we have grown were not perfect, and that the imperfection in the lattice is probably characterised either by cracks which develop along the perfect cleavage plane or by local changing of the direction of the axis in the crystal. It is important to obtain a more definite idea as to the character of these imperfections and the source from which they arise since, as will be seen from the latter part of this paper, crystals obtained by various methods of growth and which differ in flexibility and specific resistance, change their resistance differently in a magnetic field.

We first of all gave our attention to the character of the " cracks," the word being used simply in connection with the fact that the crystal breaks very easily along the cleavage plane ; by the word crack we do not necessarily mean that a slot exists in the crystal. Probably the most correct picture of a crack is that in some layers in the crystal lattice parallel to the perfect cleavage plane the atoms are disturbed from their symmetrical positions and the cohesive forces between two successive layers are reduced to a small value. When the crystal is strained these layers easily break and then real slots are formed. This view is justified by the fact that if real cracks exist in the crystal, then if the crystal is compressed perpendicular to the cleavage plane and the compressibility is measured, we ought to be able to observe a larger compressibility just before the crack closes.

I am indebted to Dr. Webster for making some experiments in this line, but he was unable to trace any larger compressibility of our crystal at the moment when it begins to be loaded.

Evidence for the existence of cracks may, as pointed out to me by Prof. G. I. Taylor, be derived from the observation of W. Spring,* that cast bismuth has a smaller density than drawn bismuth. This can be accounted for if we suppose that during solidification the multitude of individual crystals formed, having different coefficients of expansion along different axes, set up strains on cooling and open the cracks. When, however, the bismuth is drawn into a wire the strong compression force closes the cracks again. Further, cast bismuth, when placed in an electrolyte against drawn bismuth, becomes an anode. This indicates that cast bismuth has a larger energy than drawn bismuth. Prof. Taylor suggested to me that this may be due to the fact that in cast bismuth where many cracks exist there is a larger surface energy, which accounts for this excess of energy relative to drawn bismuth.

More direct evidence of the nature of the lattice imperfections was obtained by X-ray analysis.

My thanks are due to Dr. Alexander Muller, of the Davy Faraday Laboratory, for being kind enough to take X-ray photographs of various specimens of bismuth rods which had been grown by different methods, and the result of his investigation has thrown much light upon the character of the imperfections of the lattice.

First of all it was found that the crystals grown by our method are coated with a surface layer of very small crystals which are distributed at random. After etching the rod, however, and removing in all 0·1 mm. of its diameter, these small crystals disappeared and no trouble was experienced in the study of the crystal lattice. All the data given relate to such etched crystals.

Bismuth rods grown in a uniform temperature condition, as obtained in the apparatus shown on fig. 5, give very good X-ray reflections. This indicates that each rod consists of a quite good crystal with a perfect lattice, so far as can be judged from an X-ray photograph. Since we know that these rods are brittle and that cracks exist, we have to conclude that these cracks are rather widely spaced in the crystal and that the crystals between two successive cracks are perfect and have the same orientation. This view is also confirmed by the fact, already mentioned, that after breaking the crystal into small pieces, the separate pieces do not appear to be very brittle. On the other hand crystals grown with their cleavage planes perpendicular to the axis of the rod (made with the apparatus with the spiral, shown on fig. 4) and which we found to be the most flexible, showed on the X-ray photographs not only well-defined spots of reflection from the main planes, but also faint rings which have to be attributed to very small crystals differently orientated and more or less scattered in between the main lattice.

---

* ‘Bull Acad. R. Belg.,’ p. 1066 (1903): see also Desch, “Metallography,” p. 287 (1922).

Thus the difference between a flexible crystal and a very brittle one is due to the fact that in a brittle crystal the imperfections of the lattice are placed in successive layers parallel to the perfect cleavage plane, with sometimes a quite considerable distance between them, whereas in the most flexible crystals the imperfection of the lattice occurs in localised points scattered all over the crystal. This accounts for their flexibility as they have no preferable direction for cleavage. While both crystals are imperfect, in the case of the brittle crystal the imperfections are more regularly distributed and are of a rather macroscopic character.

From this we may conclude that an ideal crystal would be perfectly flexible and would have no well-defined cleavage plane.

With crystals grown on the copper plate (fig. 1) with the transverse temperature gradient and the perfect cleavage plane perpendicular to the axis of the rod, we have an intermediate case. The X-ray photograph shows rings and the crystal, as already stated, is brittle. Comparing crystals grown by our method with a crystal grown in a glass tube in the ordinary way, we find from X-ray photographs that these crystals have also enclosures with crystals distributed at random which make a definite ring on the X-ray photograph. As these rings are broken into small definite spots, we can infer that the enclosed imperfections are larger than in crystals grown by our method, indicating that our method of growing crystals with no outside strains gives a more perfect crystal lattice.

Finally, the X-ray analysis of a crystal made of impure bismuth and grown on a plate, as in apparatus 1, differs from that for a similar crystal of pure bismuth in having a more perfect lattice and showing very few traces of irregular enclosures. This confirms our observation that impure bismuth easily forms flexible crystals.

(7) *Discussion on the Source of the Cracks.*

From these experiments we may draw some conclusions as to the probable origin of the cracks.

First, since X-ray data, as well as the observations on flexibility, indicate that impure bismuth forms crystals with a good lattice more easily, we have to conclude that the imperfections in the lattice of bismuth crystals are not due to the presence of foreign atoms during the crystallisation. On the other hand we saw that the formation of the imperfections in the crystal lattice is very closely connected with the way in which the crystal is cooled and from the experiment, described in section 5, on the change of resistance during solidification, we have seen that the cracks probably occur at the moment when the crystal is solidified. We therefore have to conclude that the spoiling of the lattice is closely related to peculiarities in the crystallisation of bismuth.

I think it would be difficult to find an explanation which accords so well with the facts as the following.

We suppose that when solidified below the melting point bismuth at first forms a crystal having one crystalline modification, which, at a very slightly lower temperature, is transformed into a second crystalline modification, which we know to be the usual modification for bismuth. Since this transformation takes place when the bismuth is in a solid state and must be accompanied by changes in shape, stresses are set up in the crystal lattice which account for the spoiling of the lattice.

When the crystal is grown in a very uniform temperature (as in apparatus 5), the first modification is changed into the second one in larger sections, and between the successive sections we have an imperfect layer of the lattice which, under strain, may develop into a crack, but in between the two cracks the crystal is perfect, as has been shown already by X-ray analysis and as will be confirmed by later experiments in a magnetic field. The small specific resistance for these crystals indicates that the number of the cracks is not large.

In the case where the crystal is a flexible one and is grown with a very sharp temperature gradient perpendicular to the cleavage plane, the transition from one solid modification to the other takes place in smaller layers and in a somewhat irregular way. There will in consequence be many imperfections of the lattice scattered all over the crystal.

The view that during the solidification of bismuth, strains are set up in the crystal was confirmed by the following observation. A crystal rod was placed on a warm copper plate and a white hot platinum strip, heated by current, was placed just above it. In this way the bismuth was melted perpendicularly to the length of the rod, and then by a slow motion of the spiral along the rod, gradually solidified. In this case the cleavage plane was perpendicular to the axis of the rod. When the crystal commenced to grow it was very interesting to notice that the rod began to bend and the ends started to lift from the plate. This bending can only be accounted for by a strain set up in the crystal during the solidification. The crystal obtained in this way was extremely brittle.

If our hypothesis as to the explanation of the appearance of the cracks is correct, we have to admit a change in the shape of a bismuth crystal when it is changing from one modification to the other. We tried to confirm this by heating a bismuth rod to the melting point and observing whether a sudden change in length occurs near the melting point. Accurate experiments are difficult but indications were obtained that in some of the crystal rods there occurred a very small contraction in the length before the melting point, of the order of 0·06 per cent.

Our hypothesis is supported by some old experiments of Curie,* which confirm the existence of a second modification of bismuth slightly below the melting point. Studying the diamagnetic properties of bismuth at a temperature very close to the melting point, Curie noticed that bismuth loses its strong diamagnetic properties practically at once at a temperature slightly lower than that of the melting point, when the bismuth is still solid. This suggests that bismuth may exist in a solid state without having its strong diamagnetic properties, and this would correspond to the second modification of bismuth which is necessary for our hypothesis. From the experiments of Curie and the experiments on the change of length of bismuth during solidification, we conclude that the transition from the first modification to the second occurs at a temperature very near to the melting point, and probably only a few degrees below.

The existence of the second modification is also suggested by a phenomenon observed by Bridgman† in his pressure experiments. Usually solid bismuth, unlike other conductors, shows a positive coefficient of change in resistance when submitted to a uniform pressure. Below the freezing point, however, Bridgman observed a negative coefficient. He suggests that this change of sign is produced by strains set up in bismuth, as it is solidified in a small tube, but it is also possible to account for this difference in sign by the existence of the second modification, which is more like other metals.

---

* 'J. de Physique,' vol. 4, p. 206 (1895), footnote.

† 'Proc. Amer. Acad. Arts Sci.,' vol. 56, p. 115 (1921).

In this hypothesis the influence of the impurities which help to make the crystal flexible and more perfect may probably be accounted for by the fact that the impurities make the first modification of bismuth more unstable, and the transition takes place more gradually.

The following tentative suggestion may be made to account for the two crystal modifications. It is well known that an element of the rhombohedral bismuth crystal lattice can easily be obtained by a slight deformation of a cubic cell when it is stretched in the direction of one of the four main diagonals. On this view the first modification is a cubic one which is transformed into the second modification by this type of deformation.

The physical differences between these two crystal modifications can be explained by the idea suggested by Ehrenfest* that in the ordinary bismuth modification the strong diamagnetic properties can only be accounted for by taking the orbits of electrons inside the crystal to be of such a size that they include two centres of the lattice. In this case we have to suppose that in the first " cubic " modification the electrons move in rather smaller orbits, including only one nucleus, and that the crystal shows no strong diamagnetism, as indicated by Curie's experiments.† When the electron orbits change so as to include several nuclei, the forces of cohesion alter and the crystal is deformed and takes its rhombohedral shape and the bismuth begins to be strongly diamagnetic. Probably a close investigation of the crystal lattice by X-rays near the melting point will be necessary to give a definite proof, but at present this view is suggested as a useful working hypothesis.

In this way we can also easily explain the curious fact, observed and described in the beginning of this paper, that a slight strain set up in a crystal, when it is grown from a given crystal, will change the orientation of the perfect cleavage plane without changing the general orientation of the lattice. It is easily seen that when the crystal is growing in a " cubic " state, then this cubic modification can be transformed to the rhombohedral modification in four different ways corresponding to pulling out the elementary cubes along the four possible diagonals. The strain set up in the lattice by the cooler part of the rod, which has already transformed itself into the rhombohedral modification, will direct the transformation from the cubic state, but if we have an outside factor, such as the pressure from the walls of a tube, or the extension forces as in the Czochralski method, this factor may play a predominant part in the orientation of the trigonal axis in the rhombohedral modification. It is evident that this explanation accounts for the fact that the general orientation of the crystal is unaltered and only the perfect cleavage plane changes from one plane of the pseudo-octahedron to the other.

---

* 'Physika,' p. 388 (1925).

† *Loc. cit.*

From a closer consideration of the strains set up by the methods of growing bismuth crystals in glass tubes, or pulling them out of a crucible, we may see that there is always a tendency for the perfect cleavage plane to be orientated parallel to the main shearing stresses in the rod. The strain set up in the bismuth when it is growing in a glass tube is produced by a pressure of the walls of the tube at the moment when the bismuth is crystallised and changes its volume and shape. In the Czochralski method the strain is produced by a slight pull under the weight of the liquid part of the rod. In both cases the shearing forces are not perpendicular to the rod and probably approach an angle of 45°. This indicates why rods grown by both these methods have this preferable angle for the perfect cleavage plane.

It may also be possible that the local variation in the direction of the axis of the crystal grown with its cleavage plane parallel to its length may be due to local stressing set up in the crystal owing to the non-uniformity of the temperature gradient.

(8) *Discussion of Results.*

From the results of the previously described experiments, leaving aside the hypothesis of the origin of the imperfection of the crystals, we may explain many of the peculiar phenomena observed in bismuth, which make this metal so unlike others.

In one of the previous sections we have already given an explanation of the increase of density in drawn bismuth wires. As an instance of other phenomena it has been stated by several authors* that bismuth has two modifications, called α and β. The temperatures of transformation for these two bismuth modifications has been very much discussed and different authors have given quite different results (75°, 112° and 161° C.). From the results we have obtained we may conclude that no such α and β modifications exist, but only bismuth having more or less imperfections in its crystal structure. When, for instance, a lump of bismuth, consisting of many small crystals, is cooled, the stresses occurring between the individual crystals may open the cracks, and since the small individual crystals are probably regular in size, we may expect the cracking to occur at a more or less definite temperature. The energy absorbed in producing the cracks and the increase of volume may give the impression of the existence of two modifications of bismuth, but in another lump of bismuth, having individual crystals of another size, this cracking will take place at a different temperature.

---

* Jänecke, 'Z. f. Phys. Chem.,' vol. 90, p. 313 (1915); Cohen and Moesveld, *ibid.*, vol. 85, p. 419 (1913).

It has also been observed that bismuth has a variable temperature coefficient for its specific resistance. This again can be easily explained by the fact that in changing the temperature of the bismuth the cracks open or close and the resistance varies in a different way for each rod. In our experiments with single bismuth crystals, as shown on fig. 6, no such phenomena occurred.

Finally, we have the fact that after cooling bismuth and bringing it back to the initial temperature, the resistance is not the same. This again may be explained by the opening of cracks which failed to close when the bismuth was cooled down.

While we have mentioned here only a few cases where the existence of cracks in bismuth throws a considerable light on the peculiarities observed in this metal, we think it probable that in many other phenomena (such as observed in the Hall effect) the occurrence of cracks may offer an adequate explanation.

Unfortunately, in our experiments, we were unable to obtain an ideal monocrystalline bismuth rod, except in so far as the small pieces of bismuth rod, which broke off when a rod was obtained in the apparatus on fig. 5, were probably good crystals. From our experiments, however, we may conclude that an ideal bismuth crystal at room temperature is flexible and does not cleave at all.

It will be of interest to see how far the conclusion drawn from this investigation can be applied to other crystals. For example, is a cleavage plane in a crystal always a consequence of an imperfection in the lattice ? Do other similar crystals, such as antimony, develop their brittleness in the same way as bismuth ? Has the strain a directing influence on the orientation of the axis of other growing crystals as it has in bismuth ?

Summary of Part I.

It has been shown that during the process of the growth of bismuth crystals a small strain set up in the material has a great influence on the orientation of the trigonal axis of the crystal lattice. This strain has the effect of changing the perfect cleavage plane from being one of the pseudo-octahedral planes to one of the three remaining planes of the same pseudo-octahedron according to the character of the strain.

In order to obtain crystal rods with the perfect cleavage plane orientated in any desired direction relative to the axis of the rod, no strain must be set up in bismuth during crystallisation, and a method is described by which this is effected.

Further, it has been shown that during the process of the growth of bismuth crystals, cracks and imperfections are developed in the lattice which account for the variation of the specific resistance in bismuth observed in previous researches.

It has been shown that in the case of a perfect crystal the specific resistance along the trigonal axis is $1 \cdot 38 \,.\, 10^{-4} \pm 1$ per cent. and perpendicular to the axis $1 \cdot 07 \,.\, 10^{-4} \pm 1$ per cent. at a temperature of 16° C.

The origin of the cracks is studied and it is found that they are produced during cooling at a temperature very near to that at which bismuth is solidified. The character of these cracks depends chiefly upon the temperature gradient at the point where the bismuth is crystallised.

It has been shown that a perfect bismuth crystal at room temperature will probably have no well-defined cleavage plane and is very flexible.

The hypothesis put forward to explain the origin of the cracks is that there are two crystalline bismuth modifications; one of them at present unknown, but which is probably cubic, is transferred to the ordinary rhombohedral modification at a temperature slightly lower than the melting point. This transition is accompanied by a change of shape which accounts for the occurrence of the cracks. On the basis of this hypothesis and the phenomena stated in this paper, several physical properties of bismuth have been explained.

The Study of the Specific Resistance of Bismuth Crystals and its Change in Strong Magnetic Fields and some Allied Problems.

By P. Kapitza.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 30, 1928.)

Part II.—*The Method and Apparatus for Observing the Change of Resistance of Bismuth in Strong Magnetic Fields.*

(1) *Introduction.*

It is evident that the change of resistance of a bismuth crystal in a magnetic field, which exists for 1/100 second only and varies during the whole of this time, cannot be measured by standard methods. It is, however, possible to develop a method which gives an accuracy of the same order as that obtained in previous investigations in weaker magnetic fields produced by electromagnets, in spite of three main difficulties, viz., the briefness of the time of the experiment; the variation of the field during this time which may produce an induced e.m.f. in the crystal, and finally the small value of the resistance of the crystal, this being only a few thousandths of an ohm. The reason is that in our case we have two factors which help us out of the difficulty. The first is that the magnetic field is very strong and this increases the change of resistance making it much easier to measure. Secondly, as any resistance measurement is chiefly reduced to the absolute or relative measurement of the potential drop along the conductor when a current is passed, we made this particular drop very much larger by passing much heavier currents through the crystal than is usually permissible in ordinary experiments. This is possible because the current is sent through the crystal for one or two hundredths of a second only, and the bismuth rod cannot warm up to any great extent during this time. Another advantage of these experiments is that by using very heavy currents and large potential drops we are not troubled by any thermal e.m.fs., which in our case are relatively small, but in the usual resistance measurements with bismuth may prove very awkward. Finally, by making the experiment in one hundredth of a second we can be sure that throughout the whole of the experiment the temperature of the bismuth will remain constant.

(2) *The Method of Measuring the Resistance.*

The method adopted for the measurement of the resistance was to send a known current through the bismuth crystal, and by means of potential leads to observe the difference of potential across the bismuth crystal. Both the current and the potential difference are measured by oscillographs appropriately adjusted for the measurements. By means of these two oscillographs and a third one, which measured the current through the coil producing the magnetic field, three curves were simultaneously taken on the same photographic plate. Thus, at a given moment, the magnetic field can be deduced from the curve of the current through the coil, and the resistance of the bismuth from the two other curves. As the current was kept constant during the experiment, we can take it that the deflection of the potential oscillograph was practically always proportional to the resistance of the bismuth. On fig. 7 the general arrangement of the experiments is shown diagrammatically.

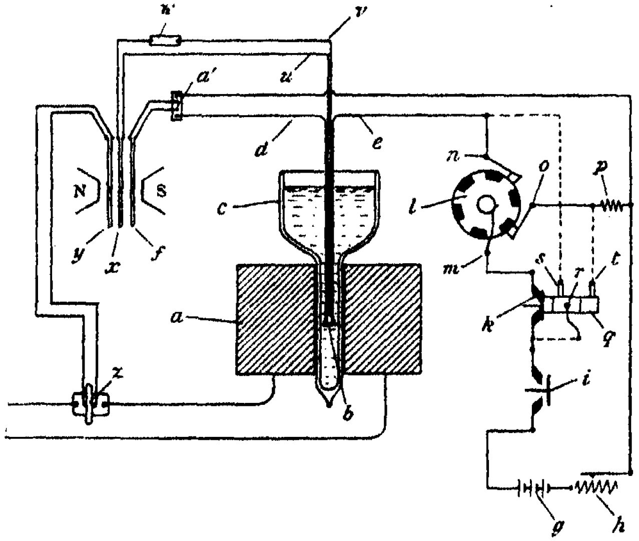


FIG. 7.

In the coil *a*, where the magnetic field is produced, the bismuth crystal rod *b* is placed. In most of the experiments the crystal was surrounded by a Dewar vessel *c*. The current is supplied to the crystal from an accumulator battery *g* through the power leads, *d* and *e*, an oscillograph *f* and a large resistance *h*. We used two alternative arrangements for sending the current through the crystal, one of them being shown by a continuous line and the other by a dotted one.

In the first arrangement, shown by the continuous line, the current was sent through a special automatically timed switch which sends the current through it for the required short time of the experiment. This switch will be described in detail later on, but its general functions are as follows. The brush *i* of this switch closes the circuit from the battery a few thousandths of a second before the magnetic field is produced in the coil. For breaking the current another independent brush *k* is used. In some of the experiments we required to make and break the current through the crystal several times during the time when the field was on. This was carried out by a very simple arrangement shown on the same fig. 7. It consisted of a copper wheel *l* with six sectors cut out and filled with insulation. Three brushes were sliding on this wheel. Brush *m* pressed on the middle of the copper disc and remained permanently in contact. Brushes, *n* and *o*, were sliding on the sectors and were brought alternately into connection with the copper wheel. The crystal and with it the oscillograph *f* were connected to the brush *n* and to one terminal of the battery. Between the other brush *o* and the terminal of the battery was connected a resistance *p* equal to that of the crystal together with the oscillograph. The third brush *m* was connected to the other terminal of the battery through the automatic switches, *i* and *k*. It is evident from this arrangement that when the disc was set in rotation impulses of current were sent through the crystal, and as the disc was worked by a little motor at 2000 r.p.m., the duration of such an impulse was about 1/400 of a second. By using two brushes, *n* and *o*, on the disc, it was possible to adjust the distance between these brushes in such a way that the current from the battery was never interrupted, and with this arrangement it was possible to make very sharp interruptions of the current in the crystal without sparking anywhere on the contacts. For the experiments in which it was not required to have such a frequent interruption of the current, the sector interrupter was left in the connections, but the disc was left stationary, in such a position that only the brush *n*, which is connected to the crystal, made contact with one of the copper segments.

The other scheme of connections, indicated by dotted lines on fig. 7, makes it possible to break the current through the crystal very sharply at any required moment during the current wave in the main coil. As will be shown in Part III, this was necessary for the experiment in which we studied the residual e.m.f. in bismuth. The principle for breaking the current used in this case is similar to that adopted in the sector interrupter just described. To the brush *k* of the automatic switch is attached a small ebonite cylinder *q* which has a copper ring *r* embedded in the middle. On the cylinder rest two brushes, *s* and *t*, made of fine copper strips (details can be seen on fig. 8 of the automatic switch). When the brush *k* of the automatic switch opens, this cylinder moves about 8 mm. along its axis, and by this motion the copper ring is shifted from the small brush *s* to the brush *t*. The brush *s* is connected, as before, to the oscillograph and the crystal, and the other brush *t*, through an equivalent resistance, to one terminal of the accumulator battery. The copper ring is connected by means of a flexible wire to the brush *i* which makes the circuit of the automatic switch. In this arrangement the brush *k* of the switch was not used, as the copper ring arrangement interrupted the current in the crystal. In this case, as in the sector interrupter, the current through the crystal was broken very sharply. It is evident that when this interrupter was used the leads to the brushes of the sector interrupter had to be disconnected and vice versa.

As will be seen from the description of the automatic switch given later, the operation of the brush *k* with the ebonite cylinder and the copper ring can be timed relative to the current wave in the coil. Thus the interruption of the current through the crystal can be adjusted to occur at any given moment in the current wave.

The potential leads, *u* and *v*, are connected through a resistance box *w* to the oscillograph *x*. There is also a third oscillograph *y* which measures the current in the coil by means of a shunt *z* in the manner described in one of the previous publications.* The three loops of the oscillographs, *f*, *x* and *y*, were placed close to each other in the air gap of the same electromagnet, and the oscillograms from them were taken on the same falling plate one above the other. The oscillographs, *y* and *f*, for the coil and for the current in the crystal were both of the same type. They had a natural frequency of about 20,000 and gave a deflection of about 3 cm. for 0·3 ampere. The current through the crystal varied from 6 to 0·5 ampere in the different experiments, and the oscillograph *f* in the crystal circuit was accordingly connected across a small shunt *a'*.

Evidently the potential oscillograph ought to be made of a different type, as only a small fraction of the power current can be used in the potential circuit. This oscillograph was made by using much finer wire for the loop and by having a greater distance between the bridges. The oscillograph actually used has a natural frequency of about 6000 per second and had a ten times greater current sensitivity than the other two elements.

---

* 'Roy. Soc. Proc.,' A, vol. 105, p. 701 (1924).

(3) *The Switch.*

The purpose of the synchronous automatic switch is clear from the previous paragraphs. It has to make and break the circuit through the power leads of the crystal at given moments relative to the current wave through the coil.

The general arrangement of the synchronous switch which satisfies these requirements is shown on fig. 8. It consists of a fly-wheel *a* which is driven by an electrical motor from the shaft *b* at a speed of 1500 revolutions per second. A standard centrifugal regulator with a contact breaker keeps the speed constant. To the fly-wheel *a* a laminated core electromagnet *c* is attached in the way shown on the drawing. This can pull up a laminated iron core *d* which rotates together with the magnet and is pressed away from the electromagnet by two small springs *e* embedded in the core of the electromagnet. By means of two brushes and two sliding rings, *f* and *g*, a current can be sent through the winding of the electromagnet, and the core pressed to the electromagnet with a large force. In the gap between the iron core *d* and the electromagnet *c* a thin iron disc *h* is placed attached to a light aluminium cylinder *i*. When the electromagnet rotates, the disc is kept in a fixed position by means of a small friction produced by the brake *k*. When the current is sent through

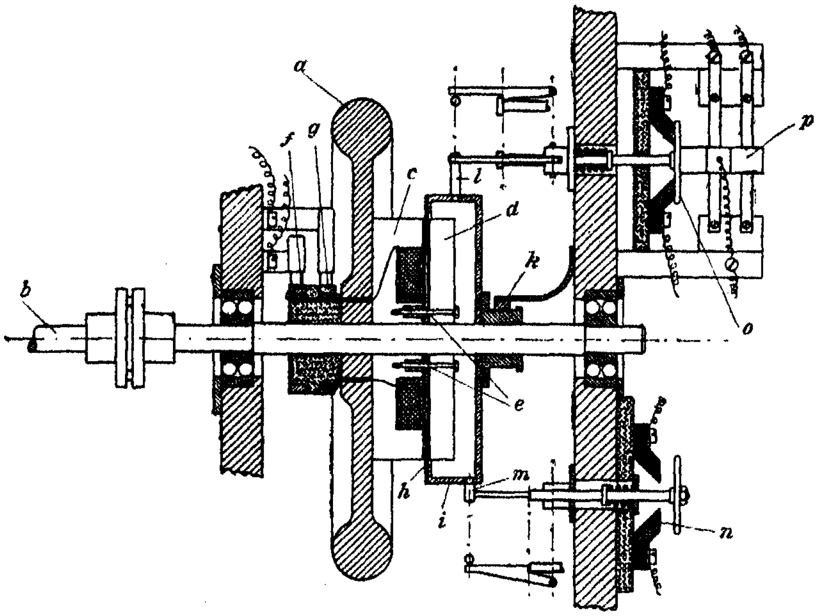


FIG. 8.

the electromagnet the core is attracted, and, owing to the large friction, the disc and cylinder begin to rotate practically at once with the same velocity as the motor. The cylinder has two adjustable radial pins, *l* and *m*, screwed in its rim. These pins, when in rotation, strike two triggers of two switches, *n* and *o*, which are released by springs and which open and close the circuits. One of these switches *o* bears the arrangement with the ring *p* described in the previous section.

During the experiment this apparatus functions in the following way. Through one of the brushes fixed to the commutator on the main switch of the big dynamo, which was described in a previous paper,* a current can be sent through the electromagnet which will set the cylinder *i*, with the pins, in rotation exactly at a definite moment before the current wave passes in the coil. By a proper choice of the position of the pins on the cylinder, these moments at which the current is made and broken can be suitably adjusted relative to the current wave. It took some time to develop the construction of this switch in order to ensure it working with sufficient reliability, the main difficulty being in making the magnetic clutch operate sufficiently quickly. By sending a strong current in the electromagnet *c* (which can be done without danger of burning the winding, as the current is sent for a few hundredths of a second only), and by making the disc *h* and the cylinder *i* light, we were finally able to adjust the moment of make and break with an accuracy of 1/5000 second. This was sufficient for these experiments.

(4) *The Fixing of the Crystal in the Coil.*

Great care has to be taken in finding the proper way of fixing the crystal in the space in the middle of the coil where the field is produced. The coil used in all these experiments had an inside diameter of 1 cm. and is described in detail in the writer's previous publication.†

To prevent the crystal from being heated by the coil, and also to make experiments at the temperature of liquid air and solid CO_2, the crystal has to be placed in a Dewar flask. This Dewar flask, shown on fig. 7, had an outside diameter of less than 1 cm. in order to fit the coil, thus leaving an inside diameter of 6 mm. only. Thus the crystal with all the leads and the holder has to be fairly small so as to fit this Dewar flask.

On the other hand, the crystal has to fit tightly in the holder as, during the experiment, electrodynamical forces try to press it sideways. These forces are relatively small—for instance, for a rod of 4 mm. length in a field of 300 kilogauss, when a current of 3 amperes is passing (these are the normal data for one of our experiments) the sideways force is only 36 grams. A much larger force due to a different cause, however, appeared in the experiments at low temperatures when the Dewar flask was filled with liquid air. In this case, when the field is on, owing to the strong paramagnetism of the liquid air, it is pulled into the Dewar flask, and the crystal with the holder is forced out. This force, for a field of 300 kg. if the cross section of the holder is 25 sq. mm., has a value lying between 1 and 5 kg., varying with the concentration of oxygen in the liquid air. The larger value of this force may be easily demonstrated by a simple experiment. The Dewar flask is placed in the coil and is filled with liquid air. A glass rod, 3 mm. in diameter and about 10 cm. long, is dropped in the flask. The field is then produced and the force just described is sufficient to throw the glass rod to a height of 7 or 8 metres. Thus it is necessary to make the holder sufficiently strong and to fasten it to a stand to avoid any displacement of the crystal.

---

* 'Roy. Soc. Proc.,' A, vol. 115, p. 671 (1927).

† 'Roy. Soc. Proc.,' A, vol. 115, p. 678 (1927).

Finally, great care has to be taken in arranging the leads to the crystal. Not only a good contact has to be established, but the leads have also to be arranged in such a way that the minimum possible e.m.f. is induced in the potential oscillograph circuit. We have to remember that in our experiments this e.m.f. may be very considerable, as the field during 1/100 of a second rises in general from 0 to 300 kilogauss and then falls again to zero. It is necessary, therefore, that the leads to the crystal shall be run bifilarly as long as possible, and the connection to the crystal arranged in such a way that no circuit is made with its plane perpendicular to the lines of force in the coil. The crystal rods have to be placed in some experiments parallel to the magnetic lines in the coil and in others perpendicular. We will first describe the perpendicular fastening which is the most difficult.

For experiments with the magnetic field transverse to the current, two methods of fastening the crystal have been generally used and both have been found satisfactory. They are shown on fig. 9, drawings A and B. In the first one (drawing A), a crystal rod *a*, about 4 to 5 mm. long and 1 mm. in diameter, is placed in a small Bakelite cylinder *b* of about the same length. This cylinder has at the ends a slightly smaller diameter, in which are made two perpendicular holes. Two silver discs *c* are fixed with shellac to the part of the cylinder with the largest diameter, and two silver wires *d*, which form the potential leads, are soldered to the rod and to the silver discs. The soldering of the silver wires to the bismuth is done by means of a spark from a condenser in the way described in Part I of this paper. As already mentioned, this does not spoil the bismuth and forms a very reliable contact.

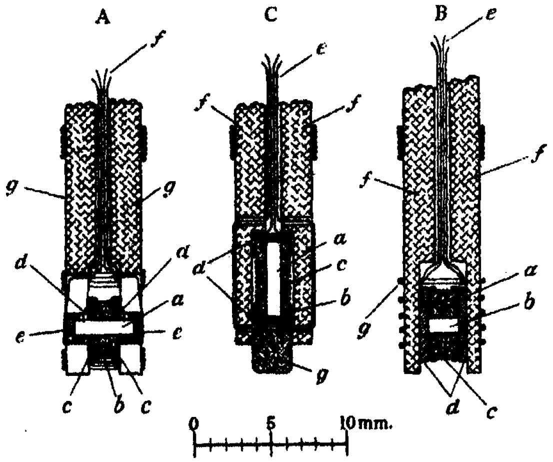


Fig. 9.

To the end of the bismuth rod are soldered two silver discs *e* by means of a low melting point solder. The silver discs form the contacts for the power leads. To make all these connections is difficult, as the dimensions are small, and it is advisable to do most of the work under a binocular low magnification power microscope.

The mounted crystal is placed in a holder. Four fine copper strips *f*, insulated from each other and pressed together between two ebonite strips *g*, which are bound together, form the holder. The end of the holder, which is connected to the crystal, is made like a double fork into which the crystal with its Bakelite fitting can be pushed. The contact between the copper strips and the silver discs is established in the following way. The two middle strips which form the potential leads each have holes at the end into which the thinner parts of the Bakelite cylinder *b* fit. They are pressed on the two silver rings and form the contact. The two remaining copper strips form the power leads and are connected to the silver disc *e* simply by making the end of the copper strips narrow, and by taking them out through the side slots of the holder and pressing them against the discs. The contacts and the steady position of the crystal are secured by a silk thread which is wound round the bottom part of the holder. It should be noticed that in such a fixing arrangement the crystal can be turned about its axis and orientated at any desired angle.

This arrangement for fastening the crystal is very elaborate and takes much time to make. It has also the disadvantage that the crystal cannot be sufficiently tightly fixed in the Bakelite cylinder, since it is only supported by the silver discs at the ends, and very often we find that when placed in the magnetic field, as a result of the electromagnetic forces, the crystal rod bends. The result is that some of the cracks, which were mentioned in the previous part of this paper, open and the resistance of the crystal increases.

This led to the development of a simple method of fixing the crystal. In this method we took advantage of the fact that the specific resistance of bismuth in normal conditions is about 75 times that of silver, and, as in our experiments we limit ourselves to an accuracy of 1 per cent., it is not necessary to make proper potential leads. It is sufficient to solder on the end of the bismuth rod two silver discs, ½ mm. thick, and to bring the power and the potential leads from these discs. Owing to the large conductivity of silver the resistance of the disc can be neglected. The arrangement for this method is shown on fig. 9, drawing B. It consists of a Bakelite cylinder *a*, cut in two parts, with a small groove in which the crystal rod *b*, about 3 mm. long, is placed. The two halves of the Bakelite cylinder are pressed together by winding round them a silk thread *c*. This provides a tight fitting of the crystal in the Bakelite, the necessity of which has just been stressed. Two silver discs *d* are soldered by a low melting point solder on to the ends of the bismuth rod, special care being taken to make a good joint with as little solder as possible. The holder consists as before, of four copper strips *e* pressed between two ebonite plates *f*. The end of the holder is made like a simple fork, and the copper strips, to which are soldered silver wire at the ends, are bent in place close to the two inside faces of the fork. When the crystal is pushed into the fork these wires touch the silver discs and form the contact. By means of a silk thread *g* wound round the fork the contacts are made safe and the crystal is kept tight. In this case, as before, the crystal rod can be turned round its axis.

The fastening of the crystal in experiments where the current was flowing longitudinally to the magnetic field is shown in fig. 9, drawing C. The crystal rod *a*, about 5 mm. long, is placed as before in a small Bakelite cylinder *b* which, in this case, is thinner and longer. The Bakelite cylinder is made in two parts and the crystal rod is fixed tightly by fastening these two parts by a thread *c*. Two silver discs *d* are soldered to the ends of the rod. The holder, as before, consists of four copper strips *e* and two ebonite plates *f*. The holder has a fork with a cylindrical opening, where the crystal with its Bakelite cylinder is inserted. To the end of the bare copper strips are soldered wires which touch the silver discs in pairs, in the way shown on the drawing. The contact between the discs and the wires is produced by a screw plug *g* which is screwed into the bottom of the holder.

In all these methods, in cases where the crystal has its cleavage plane perpendicular to the axis of the rod, special care has to be taken to prevent any tensile forces along the rod which may arise from the difference of thermal expansion of the Bakelite and the bismuth. The Bakelite cylinder, expanding more than the bismuth rod, may press the silver disc and strain the rod. This will result in opening the cracks in the bismuth and in changing its resistance. To avoid this strain a clearance has to be left between the silver disc and the Bakelite cylinder. This is easily accomplished by placing a small mica washer between the silver disc and the Bakelite cylinder during the process of soldering and then removing it.

(5) *The Method of Measurement.*

The experiment for measuring the change of resistance of bismuth crystals was performed in the following way. The required bismuth rod was fixed in the holder in the manner described in section 4 and placed in the Dewar flask, which, according to the requirements of the experiment, was either empty or filled with a freezing mixture of ether and solid CO_2 or liquid air. The absolute resistance of the crystal was measured by means of the potentiometer method described in the previous Part of this paper. The Dewar flask was then placed in the coil and the connection made according to one of the methods described in section 2. The current to be sent through the crystal was adjusted by passing it through an equivalent resistance. The resistance in series with the potential oscillograph (*w*, fig. 7) was also adjusted in such a way as to have the deflection of the oscillograph of a suitable amplitude for measurement. During the whole of this experiment the machine was used in the way described in my previous paper.* After the experiment was made an oscillograph with three lines was obtained, as shown on the reproductions. The resistance of the bismuth rod was then measured again. If the bismuth rod was properly fixed no change of resistance was observed.

After the plate was dry it was placed in an enlarging camera, which projected the oscillogram on to a piece of paper enlarged about ten times. By means of pencil the enlarged oscillogram was copied. On the same paper reference lines were traced and the deflections were measured by means of a simple mm. rule. The thickness of the lines of the oscillograms were about 2 to 3 mm.

---

* 'Roy. Soc. Proc.,' A, vol. 115, p. 181 (1927).

and the measurements were always made from the middle of one line to the middle of the zero line. The accuracy of measurement was about $\frac{1}{4}$ mm. and as the maximum deflection was adjusted to be between 50 to 80 mm., the accuracy of such a simple method of measurement was quite sufficient. From a measured deflection of the oscillogram which measured the current through the main coil, the magnetic field in the coil was determined in the way described in my previous paper.* From the two other curves the resistance of the bismuth rod in this magnetic field can be determined. As the current in the coil had a sine form, the change of the resistance of bismuth can be deduced from a single oscillogram for practically all fields from 0 to the maximum value obtained in the experiment. Actually, in most of the experiments, the potential curve only is used, as the current through the crystal remains constant and has a value adjusted beforehand. Thus the ordinates of the potential oscillograph curve are proportional to the resistance of the bismuth.

The relative change of resistance is obtained by dividing the deflection of the potential oscillograph, measured with the field, by its deflection without the field. This deflection can either be calculated from the measured resistance of the crystal and the constants of the oscillograph and of the circuit, or measured from the oscillogram directly. This, however, can only be done on oscillograms taken in weak fields where the resistance of bismuth does not change more than ten times, as, for instance, in oscillogram No. 5, Plate 6. On this oscillogram the current started before the field was established, and the initial deflection was easily measured. But when the change of resistance is about 100 or more times, and the resistance in series with the potential oscillograph is very large, the initial deflection is so small that it cannot be measured correctly (*see* oscillogram No. 8, Plate 7). It must be mentioned that such a simple calculation for the change of resistance of bismuth can only be used when the change of resistance is small—up to 50 or 100 times, otherwise the increased resistance changes the condition of the circuit appreciably. For instance, if we have an increase of resistance of 1000 times or more, as happens at the temperature of liquid air, the resistance of the crystal rod, which is initially between 0·002 to 0·005 ohm, will increase to a value of 2 to 5 ohms. In this experiment only half an ampere can be safely sent through the crystal in order to avoid heating it, even in 1/100 second. The potential drop across the crystal will be 2 to 5 volts, and if an accumulator battery of 30 volts is used the main current through the crystal will diminish by 6 or 7 per cent., and a correction has accordingly to be introduced.

---

* *Loc. cit.*, p. 679.

The potential oscillograph consumed about 15 milliamps. for its full deflection. When a current of 3 amperes was passing through the crystal, the measurement was not affected by more than ½ per cent. But in the cases where the current through the crystal was limited by heating considerations to one ampere or less, the current through the potential oscillograph is already comparable with the current through the crystal, and a correction has to be introduced. The introduction of these corrections is simple and was always made at any moment when it was more than 1 per cent. No special mention of this subject will be made later on in the paper, and in all subsequent discussions it will be assumed that these corrections have been made. When the resistance of bismuth changes more than 2000 times, the correction commences to be rather large and the measurement becomes less reliable. Indeed, with such an increased resistance, bismuth has a specific resistance 200,000 times that of silver and is more like a non-conductor. Special methods of accurately measuring the resistance beyond this point are quite possible but have yet to be worked out.

(6) *The Errors and Stray Effects.*

The sources of error have to be considered rather carefully in our experiments, as in any new method of experimenting. We have already mentioned that we limited the accuracy of the experiments to 1 or 2 per cent., this being quite sufficient in this kind of experiment, as it will be seen that the error which arises from the unknown purity of bismuth may be considerably larger.

From the method of measuring the oscillograms, and by carefully calibrating the oscillograph, we are fairly sure that the quantity measured from the oscillograms represents the actual values well within 1 per cent. All the possible errors occurring in this method of measurement with oscillographs have already been fully discussed in the author's previous papers.*

The next source of error is the e.m.f. in the potential leads, which is induced by the variation of the magnetic field during the experiments. Experiment shows that by the arrangement for fixing the crystal, described in section 4, where all the leads have been arranged bifilarly, the induced e.m.f. is made very small compared with the potential drop along the crystal, and the effect produced in most cases is negligible within the limits of experimental error. This is easily proved by taking an oscillogram with the battery (*g*, fig. 7) disconnected. The deflection of the potential oscillograph is then entirely due to this effect. On fig. No. 1, Plate 5, such an oscillogram is reproduced. In this case the amplitude of the magnetic field reaches the value of 320 kilogauss, and only by very careful examination can we detect a slight curving in the line traced by the potential oscillograph.

---

* 'Roy. Soc. Proc.,' A, vol. 105, p. 701 (1924).

In certain experiments where the crystal was not fastened so successfully, and in cases where the resistance in series with the oscillograph was small, we were able to observe at the beginning and the end of the current wave a slightly larger deflection. In these cases the error can be easily eliminated by taking two oscillograms with the same field, but with the direction of the current through the crystal reversed. Taking the average of the measurements from the two oscillograms, the disturbing e.m.f. will be eliminated. This e.m.f. was also automatically eliminated in all the experiments when we used the sector interrupter, which was described in section 2. With the sector interrupter we obtained oscillograms similar to Nos. 5 to 10 (Plates 6 and 7). The necessity for using the sector interrupter arose in cases where it was necessary to eliminate the e.m.f., which occurs in bismuth crystals in a magnetic field when a current is passing through the crystal. By use of the interrupter this e.m.f., and the induced one are automatically eliminated. In these oscillograms, obtained with the sector interrupter, a new source of error may come in as the potential oscillograph now works under severe conditions. During two-hundredths of a second it has to swing across the plate over ten times from below zero to the full deflection. As the frequency of the potential oscillograph was comparatively low (6000), and comparable with the time of a deflection, the time lag of the deflection is noticeable on all the oscillograms taken. It can also be noticed that the oscillograph was slightly under-damped. As we required to measure the difference of the negative and positive deflections at the moment when the current is broken or started, an extrapolation of the deflection was necessary which naturally made the measurement less accurate. We estimate the error of this extrapolation in the worst cases to be about 3 per cent.

Finally, we have to consider the effect of the heating of the crystal produced by the current. It becomes serious only when we deal with the crystal at the temperature of liquid air and in strong fields. The specific resistance of bismuth is in this case so high that even a small current may heat it up. Let us make a numerical examination. The heat generated in the crystal is

$$Q = 0 \cdot 24 \int_0^T (Is)^2 \, R dt, \tag{1}$$

where I is the current density in the crystal, s is the cross section and R the resistance. With sufficient approximation (as will be seen from the next Part of this paper) we may take R to be proportional to the magnetic field, and, as the magnetic field changes sinusoidally, we may take

$$R = R_{\text{max.}} \sin (\pi t/T), \tag{2}$$

where T is the half period of the current wave. We obtain from (1)

$$Q = 0 \cdot 15 \, R_{\text{max.}} \, T \, (Is)^2. \tag{3}$$

It is easy to see that the rise of temperature will be

$$\delta t = 0 \cdot 15 n \sigma I^2 T / kd. \tag{4}$$

In this formula $n = R_{\text{max.}}/R_0$ the relative increase of resistance; k and σ are the specific heat and resistance, and d the density. At liquid air temperatures $k = 0 \cdot 028$, $\sigma = 0 \cdot 5 \,.\, 10^{-4}$, $d = 9 \cdot 80$ and in our experiments $T = 1/40$, so that we get

$$\delta t = 0 \cdot 68 \,.\, 10^{-6} \, n I^2.$$

Now in our experiments with a bismuth rod of 1 mm. diameter, if we take the current $i = 0 \cdot 5$ amp., and the increase of resistance is about 2000 times, we get an increase of temperature $\delta t = 5 \cdot 5°$, which is already appreciable.* On the other hand we cannot greatly decrease the current, as then the correction for the potential oscillograph will be too large.

This, however, only happens in liquid air experiments; in other experiments the rise of temperature is never greater than $0 \cdot 5°$.

These are practically all the errors which can be introduced into the experiments by this method of measuring the resistance change in the magnetic field. The other errors which arise from the impurities of bismuth, from the imperfection of orientation of the crystal and similar causes, will be considered later in connection with each individual experiment.

Summary.

A description is given of a method for measuring the change of resistance of a conductor when it is placed in a magnetic field which exists for 1/100 second.

The study of the change of resistance is made possible by the fact that larger current densities are permitted in the conductor, for such a short time, without heating the conductor, and no specially sensitive apparatus is required. The measurements can be made with an oscillograph.

---

* When the sector interrupter is used the increase of temperature is about half.

A special switch is described which permits the current to be sent through the investigated conductor for a very short time, during the existence of the magnetic field. The make and break of this switch may be adjusted relative to the current wave through the coil, which produced the magnetic field with an accuracy of about 1/5000 second.

The experimental arrangements are described which permit the study of the change of the resistance of bismuth crystals in a magnetic field at different temperatures.

The elimination of errors, due to the induction effects and heating up of the crystal during the process of measurement, are discussed.

The Study of the Specific Resistance of Bismuth Crystals and its Change in Strong Magnetic Fields and some Allied Problems.

By P. Kapitza.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 30, 1928.)

Part III.—*The Change of Resistance of Bismuth and the Time Lag in Magnetic Fields.*

Introduction.

The study of the change of resistance of bismuth crystals in a magnetic field is an involved problem in which a number of variables have to be taken into account.

In the first place the change of resistance of bismuth crystals depends on the strength of the magnetic field; secondly, on the direction of the current relative to the axis of the crystal; thirdly, on the orientation of the crystal in the magnetic field, and finally, on the temperature. It also depends on the chemical purity of the bismuth, on the state of perfection of the crystal lattice and on the strains which may be set up in the crystal.

In order to study these phenomena a great deal of experimental work is required, the difficulty being increased by the time lag phenomenon. A large amount of generalisation can be achieved, however, if we combine the results with some of the phenomena described in Part I of this paper, especially with the observations on the development of imperfections and cracks in the bismuth crystal.

(1) *The E.M.F. which accounts for the Time Lag and its Elimination in the Resistance Measurement.*

Before commencing the measurement of the change of resistance of bismuth crystals, it was necessary to eliminate the time lag phenomena and to determine the true resistance of bismuth in magnetic fields. The magnitude of the time lag can be very easily seen on oscillogram No. 2, Plate 5, for a crystal where the current is passing perpendicular to the cleavage plane, and where the current is also perpendicular to the lines of force of the magnetic field; for this orientation, the lag phenomenon is most marked. We see from the curve p traced by the potential oscillograph, that during the time from x to y, when the current in the coil is constant and the magnetic field also remains constant, the deflection changes its amplitude quite appreciably, beginning with a high value and then gradually decreasing. As the deflection of the potential oscillograph is proportional to the resistance of the crystal, this decrease means that the resistance of bismuth is diminishing with the time. The cause of this diminution can easily be seen on the same oscillogram. In the second half of the current wave, when the current through the crystal is suddenly broken, the deflection of the potential oscillograph does not drop to zero, but retains a quite definite negative value. This means that along the crystal there exists a residual e.m.f. when there is no current. This residual e.m.f. in the crystal reduces the fall of potential due to the ohmic resistance as it is in the direction "to help" the current to pass through the crystal. The origin of this e.m.f. will be considered later in a special paragraph, but at present it suffices to mention that it has already been traced by Seidler,[*] and has been studied by several investigators.

We shall now describe the method by which this e.m.f. in the crystal can be eliminated and the actual potential drop along the crystal due to the change in the ohmic resistance obtained. In the first place we see from the fact that the deflection of the potential oscillograph diminishes during the time when the field is constant, that the e.m.f. gradually increases in value with the time. This can be seen in oscillograms (Nos. 2 and 3, Plate 5) for the same crystal, under the same conditions, except that the current was broken, by means of the switch described in Part II, section 3, at two different moments during the flat part of the current wave, in (3) near the beginning, and in (2) towards the end. It is easily seen from the negative deflections of the potential oscillograph that the e.m.f. is larger when the current is broken later. On these two oscillograms we take the full swing of the potential oscillograph, at the moment when the current through the crystal is broken, to be proportional to the ohmic potential drop along the crystal, since the deflection before the break is diminished by the amount of the residual e.m.f. This assumption is confirmed by direct measurements of these oscillograms (Nos. 2 and 3), which show that the amplitude of swing is independent of the moment at which the current is broken. In this way, therefore, we are able to eliminate the time lag phenomenon from our experiments. As the development of this residual e.m.f. takes some time, it appears that by this method we measure the resistance which would occur in an ideal experiment in which the current is suddenly established in the crystal and in which the e.m.f. has no time to develop. Later on, when we consider the origin of this e.m.f. in section 6, we shall give a more complete justification [illegible] this method of measuring the resistance.

---

[*] 'Ann. d. Physik,' vol. 32, p. 337 (1910).

The elimination of the e.m.f. for several points along the current wave is obtained by using the sector interrupter, which has been described in Part II, section 2, of the present paper. Oscillogram No. 4, Plate 5, is taken for a crystal without the sector interrupter. If we plot the curve of the relative change of the resistance of the crystal taken from this oscillogram, we obtain a curve in the shape of a loop, as shown on fig. 10, by means of a broken line. It is obvious that this loop results from the asymmetrical form of the deflection of the potential oscillograph, which is due to the e.m.f. developed in the crystal.

To eliminate the e.m.f. in the crystal we take another oscillogram similar to oscillograms Nos. 5, 6, 7 and 8, Plates 6 and 7, in which the sector interrupter is used and take the true ohmic potential drop to be equal to the full swing at the beginning and at the end of each interruption, in the way described for oscillograms Nos. 2 or 3. The curve obtained in this way is shown on fig. 10 by means of a continuous line, and it is seen that the points taken on the side of the increasing fields, marked by circles, and the points taken on the descending part of the current wave, marked by crosses, lie on a practically continuous line within the limits of experimental error. As already mentioned in this method of elimination of the e.m.f. by means of the sector interrupter, the lag in the deflection of the oscillograph begins to be appreciable and the accuracy of the experiment is diminished.

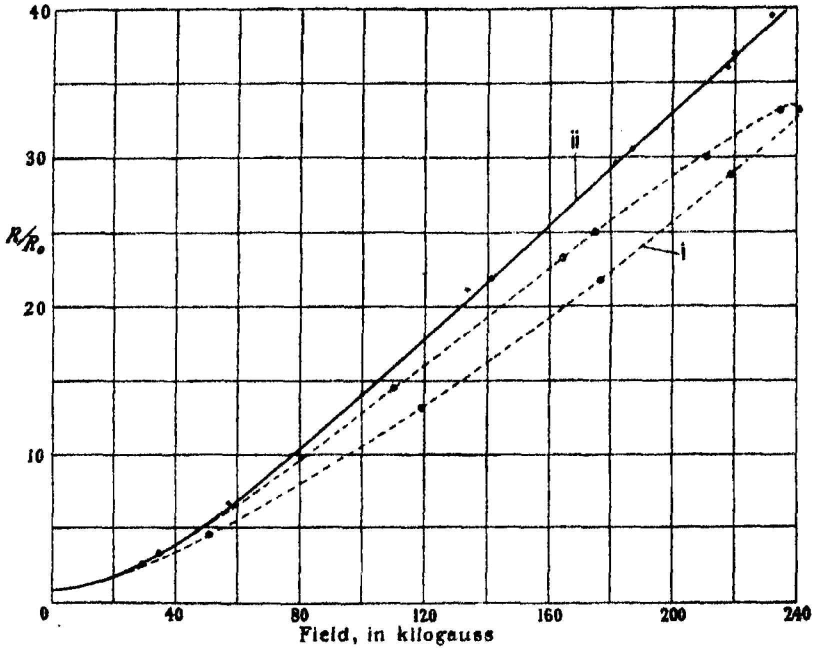


FIG. 10.—Crystal orientated I ‖ Axis and I ⊥ H temperature 17° C. Curve i taken from oscillogram No. 4 without using the sector interrupter and the residual e.m.f. is not eliminated. Curve ii taken with the sector interrupter and the residual e.m.f. is eliminated.

(2) *The Change of Resistance of Bismuth Crystals in a Magnetic Field with the Current perpendicular to the Field.*

Some of the results of our experiments are represented by curves on diagrams, figs. 11 to 16, the ordinates representing the ratio of the resistance R of the crystal in a magnetic field to its resistance R_0 without the field, and the abscissæ the magnetic field in kilogauss. Curves marked by the same letter refer to the same crystal, the suffix indicating the orientation of the crystal relative to the magnetic field.

Most of the curves range from 0 to 320 kilogauss and each was obtained from at least two oscillograms. The first was taken with weaker fields and covered the range from 0 to 100 kilogauss, one or more additional oscillograms covering the remaining part of the range to 320 kilogauss. On some of the curves the measured points are marked to give an idea of the accuracy of the experiments; on others this has not been done for the sake of clearness in the drawing. For illustration some oscillograms are reproduced, the details of which can be learnt from the description.

The experiments for most of the crystals were made at three different temperatures. First at room temperature, secondly at a temperature of about − 80° C., this being obtained by filling the Dewar flask *c* of fig. 7, in which the crystal was placed, with a mixture of solid CO_2 and ether, and finally at a temperature of −182° C. by filling the Dewar flask with liquid air.

We will first describe the change of resistance of bismuth crystals when the current is flowing along the trigonal axis, *i.e.*, perpendicular to the cleavage plane. For abbreviation we call this case I $\perp$ H and I $\parallel$ Axis, and the corresponding experimental results are represented in diagrams, figs. 11, 12 and 13.

We first consider the effect of the orientation of the crystal when it is rotated round the trigonal axis, six similar orientations of the crystal relative to the direction of the magnetic field being possible in this case. In the first set one of the lines on the perfect cleavage plane which, as was stated in Part I of this paper, corresponded to the intersection of one of the three remaining pseudo-octahedral planes with the perfect cleavage plane, is perpendicular to the lines of magnetic force, and the other two lines form an angle of 30° with the magnetic field. (This we shall call Line $\perp$ H.) In this case we have the maximum change of resistance of bismuth in a magnetic field (given by curve A_1). When the crystal is turned through 30° two of the three lines on the perfect cleavage plane form an angle of 60° with the magnetic field, the remaining line being parallel to the field (Line $\parallel$ H) and the increase of resistance with the strength of the magnetic field is a minimum (curve A_2).

I am indebted to Dr. Webster who studied this phenomenon in weak magnetic fields at room temperature, and who found that in turning a crystal round its axis in a magnetic field, these minima and maxima repeat themselves with a periodicity of 60°, and the variation in the change of resistance in a constant field almost approaches a sine curve. This effect of orientation is illustrated in oscillograms 5 and 6, taken for the same crystal under exactly similar conditions, except that in oscillogram 5 the crystal was in the position to give the maximum change of resistance (Line $\perp$ H) whilst in oscillogram 6 the crystal was turned round its axis by 30° (Line $\parallel$ H) and the minimum change of resistance was obtained. On comparing the two oscillograms, the smaller amplitude of the potential oscillograph traced by curve *p* is evident. In curves figs. 11, 12 and 13, where the change of resistance for the three above-mentioned temperatures is represented, the crystals with the orientation which gives the maximum change of resistance are marked by the suffix 1, and correspondingly the crystals in the minimum position bear the suffix 2.

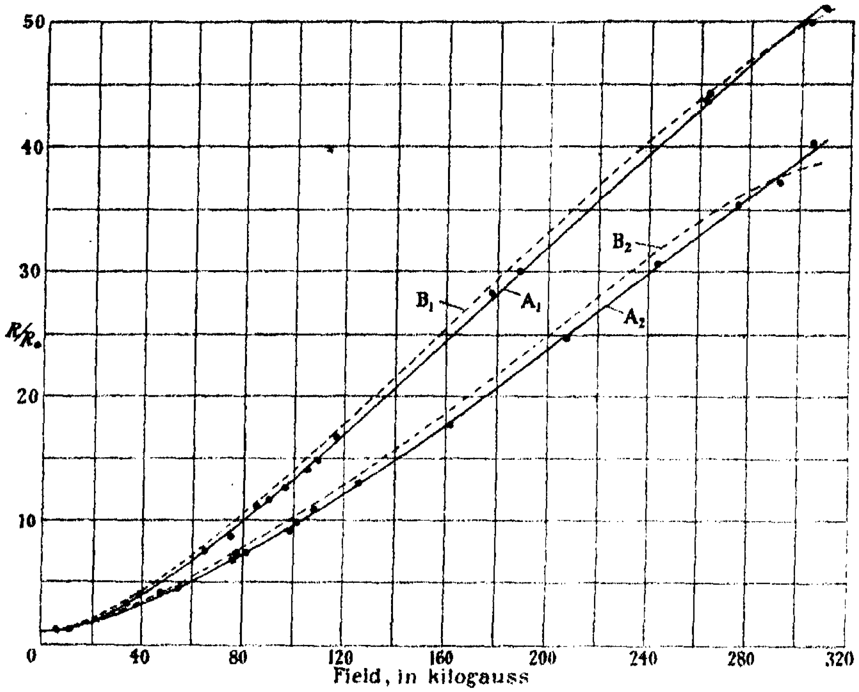


FIG. 11.—Room temperature, orientation I || Axis and I ⊥ H. Curve A_1, Crystal A, Line ⊥ H. Curve A_2, Crystal A, Line || H. Curve B_1, Crystal B, Line ⊥ H. Curve B_2, Crystal B, Line || H.

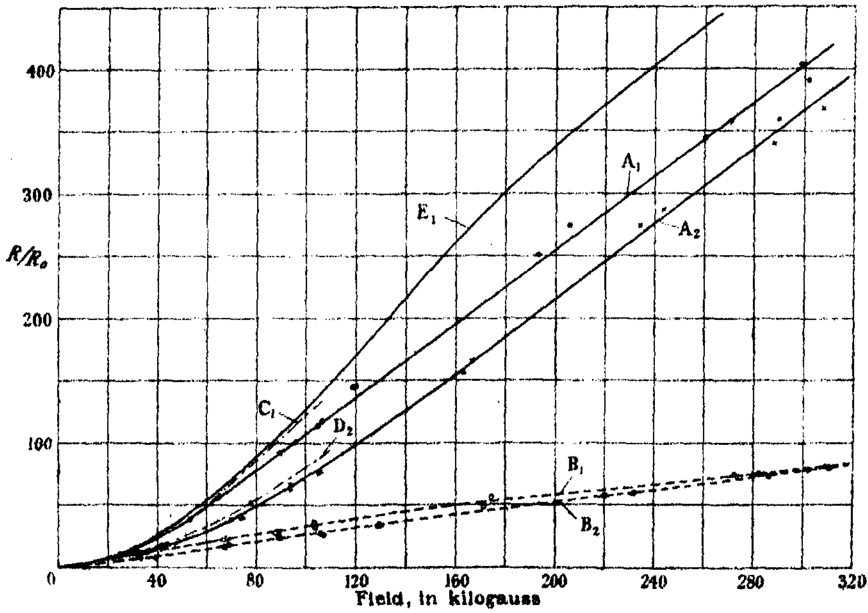


FIG. 12.—Temperature of the mixture of solid CO_2 and Ether, Orientation I || Axis and I ⊥ H. Curves A_1, B_1, C_1, E_1, Line ⊥ H. Curves A_2, B_2, D_2, Line || H.

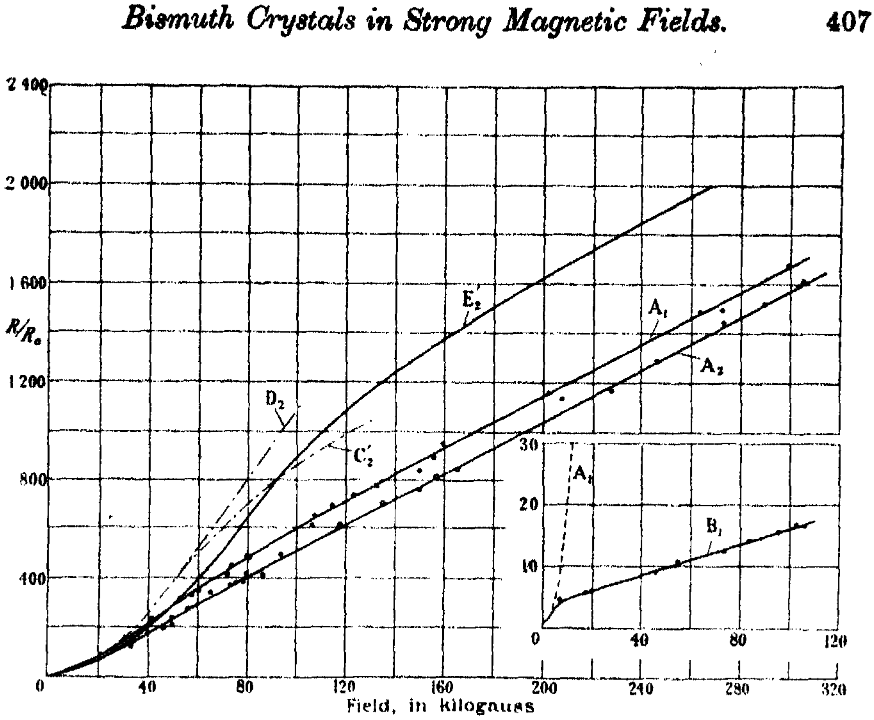


FIG. 13.—Temperature of Liquid Air, Orientation I ∥ Axis and I ⊥ H. Curves A_1, B_1, Line I ⊥ H. Curves A_2, B_2, C'_2, D_2, E'_2, Line ∥ H.

The effect of the orientation has been most completely studied in the case of crystal A. We notice that the difference between the ordinates of the curves A_1 and A_2 increases up to a certain field and then remains practically constant, the curves being very nearly parallel or even showing a slight tendency to coincide at very strong fields. It is also seen that the constant difference is reached at smaller fields when the temperature is lower. It is reached in liquid air at about 70 kilogauss, in solid CO_2 at about 120 kilogauss, and in air the parallelism is not completely attained even at 320 kilogauss.

It is well known that the effect of the temperature on the resistance change is very great; it is illustrated by the three oscillograms 5, 7 and 8, taken for the same crystal A, in the same orientation, but for three different temperatures. Oscillogram 5 was taken from the crystal at room temperature. The crystal had an initial resistance of 0·00325 ohm (12° C.) and the deflection of the potential oscillograph (curve *p*) is very noticeable, even without the field. Oscillogram 7 corresponds to a temperature of —80°. The initial resistance of the crystal was 0·00215 ohm, and the sensitivity of the potential oscillograph was diminished four times (by diminution of the current through the crystal and increasing the resistance, *w*, fig. 7), the initial deflection without the field being scarcely noticeable. Oscillogram 8 is for liquid air. In this experiment the magnetic field was about three times larger than in the two previous cases, and the potential oscillograph was made 36 times less sensitive. The initial deflection of the oscillograph can no longer be traced on curve p.

The general way in which the resistance alters with the magnetic field at different temperatures can be seen from fig. 11, where the results are given for experiments at room temperature, from fig. 12 for a temperature of solid CO_2 and ether, and from fig. 13 for liquid air. We shall discuss these results in detail later, but at present we may remark that all the curves are of a similar character. At the beginning the resistance changes as the square of the magnetic field, and at stronger fields the change is proportional to the field.

We will now consider the influence of chemical impurities in the bismuth. For this purpose we compare two crystals, A and B.

The crystal A, for which the curves A_1 and A_2 corresponding to the two orientations were just described, was grown from a Hartmann and Braun rod on the copper plate (fig. 1), with the cover B, and had, on measurement, a specific resistance of $1 \cdot 51 \,.\, 10^{-4}$. The cleavage plane formed an angle of 89° with the axis of the rod. The crystal was fixed in the holder, fig. 9, B.

The second crystal rod B was grown in a similar way, but was made from ordinary bismuth obtained from Kahlbaum. The crystal was fastened in the holder in the same way as crystal A. The change of resistance in the magnetic field for this crystal is given, in figs. 11, 12 and 13, by curves B_1 and B_2. In fig. 11, at room temperature, it is to be noticed that curves B_1 and A_1 and B_2 and A_2 agree well within the limits of experimental error. At a temperature of solid CO_2, however, the curves for the crystal B, made of Kahlbaum bismuth, show quite a different change of resistance from that for crystal A. At first, up to about 15 kilogauss, they practically agree with curves A_1 and A_2, but after this field it is seen that the curves B_1 and B_2 are well below the curves of A_1 and A_2, corresponding to a much smaller change of resistance. We see from fig. 13 that at a temperature of liquid air the Kahlbaum bismuth rod gives a very small change of resistance. The results are shown on the right-hand bottom corner of the same drawing. The change of resistance of crystal B was so small, in comparison with crystal A, that it was necessary for the scale to be considerably increased in order to draw the curve. It is seen that up to a field of 8 kilogauss the Kahlbaum bismuth rod has the same change of resistance as the Hartmann and Braun bismuth rod, which is shown by a dotted line A_1 on this figure. Afterwards there is a sharp bend and the curve then runs in practically a straight line and the increase of resistance is very small.

These experiments confirm the already known fact that impurities have a large effect in diminishing the change of resistance of bismuth in a magnetic field, and we see that this is chiefly noticed in strong fields and at low temperatures. As a practical result in our case we deduced that Hartmann and Braun bismuth was of a higher chemical purity than Kahlbaum bismuth. This fact was made use of in judging the purity of bismuth for our experiments, described in Part I, when we discussed the influence of impurities on the growth of crystals.

We shall now describe one of the most interesting points, namely, that the imperfection of the crystal lattice also has an important effect on the change of resistance. In the First Part of this paper the various imperfections of the lattice were described, and it was shown how these imperfections depend upon the manner in which the crystal is grown. We have now investigated the effect of different methods of crystal growth on the change of resistance.

We first chose two samples of crystals made of Hartmann and Braun bismuth, both grown with a sharp temperature gradient along the rod, as obtained in the apparatus with the platinum spiral (fig. 4). One of these (D) was the most flexible crystal we had ever been able to obtain, having its cleavage plane inclined at 75° to the axis of the rod. The second crystal (C) was less flexible, and had its perfect cleavage plane at an angle of 89° to the axis of the rod. Both these crystals had the normal specific resistance, viz., $1 \cdot 365 \,.\, 10^{-4}$ and $1 \cdot 38 \,.\, 10^{-4}$.*

At room temperature the change of resistance of these two crystals was exactly similar to that shown by crystal A and represented by curves A_1 and A_2 on fig. 11. We have not drawn these two curves simply because they could not be distinguished from each other. At a lower temperature, however, in solid CO_2 and ether, the change of resistance of the flexible crystals was larger in stronger magnetic fields. The very flexible crystal is shown by curve D_2 and the less flexible one is represented by curve C_1. We see that in both these cases the resistance changes are practically the same as for the crystal A for fields up to 60 kilogauss, but for higher field strengths the crystals C and D give a larger increase of resistance.

At the temperature of liquid air we have even more marked differences between curves for the flexible crystals D and C and the crystal A, as can be seen from fig. 13. Here again, however, the difference appears only for stronger magnetic fields, and at a field of 100 kilogauss the flexible crystal C (curve C_1) gives a change of resistance which is about twice greater than that obtained for the crystal shown by curve A_1.

---

* The crystal D had the lower resistance, due to the inclination of the cleavage plane. This variation of resistance is in agreement with the Voigt-Thomson law ($\cos^2 \theta$) which we found to hold very well for our crystals.

This difference in the behaviour of crystals indicated clearly that it was due to variations in the perfection of the crystalline lattice, and this inspired the X-ray investigations of the crystals described in the First Part of this paper. As we have seen in Part I, the result of these investigations was to show that neither the crystal A grown on the plate, nor the flexible crystals C and D, grown in the apparatus of fig. 4 with a platinum spiral, have a perfect crystal lattice. It was also seen that the relatively best crystals were obtained by growth under very uniform temperature conditions as produced in the apparatus of fig. 5. We have chosen one of the best of the crystal rods grown in this apparatus, which had a specific resistance $1 \cdot 38 . 10^{-4}$, the cleavage plane having an angle of 89° with the axis of the rod, and studied the change of resistance in the magnetic field. We denote the curves for this crystal by E. This crystal was very carefully etched in order to remove the surface layer, which, as had been seen from X-ray examination, might contain small crystals. It was placed in a holder similar to the one given on fig. 9, B, only in this case special care was taken in handling not to produce any strain in the crystal when it was fixed. Furthermore, to avoid any strain from the holder, the two ebonite plates of the holder *f*, which pressed on the sides of the crystal, were removed and the wires were soldered to the silver discs *d*.

No difference in the change of resistance for this crystal E from that of crystals A and B was observed at an ordinary temperature, and we obtained a curve which again coincided with the curve A_2 given on fig. 11. At the temperature of CO_2, the difference between the curves E_1 and A_1 was much more marked. The change of resistance in E is larger and is straight only after 200 kilogauss.

Before describing the results for the change of resistance of crystal E at the temperature of liquid air, mention should be made of certain difficulties in the exact measurement of the ratio of R/R_0, which were met with in all the crystals at such low temperatures. The first of these difficulties is that the very brittle crystals when immersed in liquid air become very fragile and often break by themselves along the cleavage planes. This is a result of the opening of one of the cracks in the rod which is difficult to avoid in this kind of crystal (*see* Part I). As this happened with the crystal E, in order to study it at liquid air temperature we had to take another bit from the same crystal rod, and, before immersing this in liquid air, bound it with a silk thread.

This prevented the crystal from cracking. In order to distinguish this new piece we called it E′. Another difficulty is the breaking of the solder by which the silver discs were fastened to the rod. This was probably due to the strains set up between the bismuth and the silver due to a difference of expansion coefficient, and to the fact that at the temperature of liquid air the solder is very fragile. In some cases the disc did not break completely away, but the resistance of the crystal after being submitted to low temperatures increased by 10 or 20 per cent. This was probably due to cracks developed in the solder, as after resoldering the crystal the resistance generally returned to the initial value. In some cases, when the crystal was imperfect, the resistance of the rod retained its larger value, this being due to the imperfections developed in the crystal. Before each experiment and after, we measured the resistance of the crystal after bringing it to room temperature. If a variation of more than 5 per cent. was noticed, the experiment was not taken into account.

Finally, as will be seen from the next paragraph, for a given magnetic field the resistance changes approximately as the fourth power of the temperature, and at liquid air temperature a difference of 1 degree will result in a change of about 5 per cent. in the absolute value of the increase of resistance. We may easily admit that between individual experiments a difference of 1 or 2 degrees occurred from the variations in the temperature of the liquid air, which result from a difference in the concentration of oxygen.

All these sources of error probably reduce the accuracy of the absolute value of R/R_0 taken at liquid air temperature for different samples of crystals. The error in some cases may reach a value of 15 to 20 per cent. The general shape of the curve, however, is not affected by this error, which is too small to account for the large differences which were observed between the crystals A, C and D, as shown by the curves on fig. 13.

The same considerations apply to the curve E'_2 in the same diagram for the crystal E taken at the temperature of liquid air. We see that the change of resistance is again about twice as large as for crystal A, and the curve has a different shape, becoming a straight line only at the higher field of about 140 kilogauss.

The reason why we did not wish to place the crystal E in an ordinary holder, where it would be slightly compressed by the ebonite fork, was that we observed that a strain set up by the longitudinal pressure in a crystal influences the change of resistance of bismuth in magnetic fields. This phenomenon, which only happens at the temperature of liquid air, was accidentally traced from the observation that if a crystal is carelessly soldered to the silver discs, without a mica washer, as described in the Second Part of this paper, a different change of resistance is obtained. In order to prove the existence of this effect we soldered a bismuth crystal between two silver plates and measured the change of resistance in liquid air. Then, by means of a silk thread, we bound the crystal as strongly as was possible without spoiling it, and again measured the change of resistance. We found then that in strong fields the crystal gives a change of resistance which is about 10 per cent. less than that obtained for crystal which is not so bound. At the temperature of solid CO_2, or room temperature, this effect was not observed. This indicates that a pressure set up in a crystal has a similar effect to imperfections of the lattice in diminishing the change of resistance in magnetic fields.

We shall now describe the results for other orientations of the crystal, taking the case when the cleavage plane is parallel to the axis of the rod and the current is passing perpendicular to the magnetic field. We will call I ⊥ H and I ⊥ axis. In this case we have to consider the effect of the orientation of the cleavage plane relative to the field and study two main cases—first where the cleavage plane is parallel to the lines of magnetic force and the second when it is perpendicular. These two cases we will call respectively axis ⊥ H and axis || H. Two further cases ought to have been considered, depending upon the position of the lines on the cleavage plane relative to the direction of the current, but as it is impossible to alter the direction of the current relative to these lines by a mere rotation of the crystal, it would be necessary to regrow the crystal with a different orientation of the lines. We should then be faced with a possible difference in the perfection of the two crystals which at present we cannot estimate in an independent way and which might completely vitiate the comparison. We shall not therefore consider these two cases.

For the crystals with the cleavage plane parallel to the axis of the rod, we have not made that systematic study of the effect of impurities and imperfections of the crystal lattice on the change of resistance which was attempted for the perpendicular orientation. We simply took a crystal, which we call X grown from Hartmann and Braun bismuth, under the most uniform temperature conditions in the apparatus of fig. 5, this crystal having a specific resistance of $1 \cdot 07 . 10^{-4}$ at 14° C., which is the smallest we have observed. From the low value of the specific resistance, we assumed that this was the most perfect crystal which we were able to obtain. It happened that in this crystal two of the lines on the perfect cleavage plane, resulting from the intersection of the three remaining planes of the pseudo-octahedron, formed an angle of nearly 60° with the direction of the rod. For a piece cut from this crystal for each of the three temperatures which we had been using, we took two curves of change of resistance, having the crystal orientated with the axis parallel and perpendicular to the magnetic field. These two orientations are obtained by turning the crystal rod round its axis in the holder A or B on fig. 9. In the first position (axis $\perp$ H) the change of resistance in the magnetic field is larger and is given on figs. 14, 15 and 16

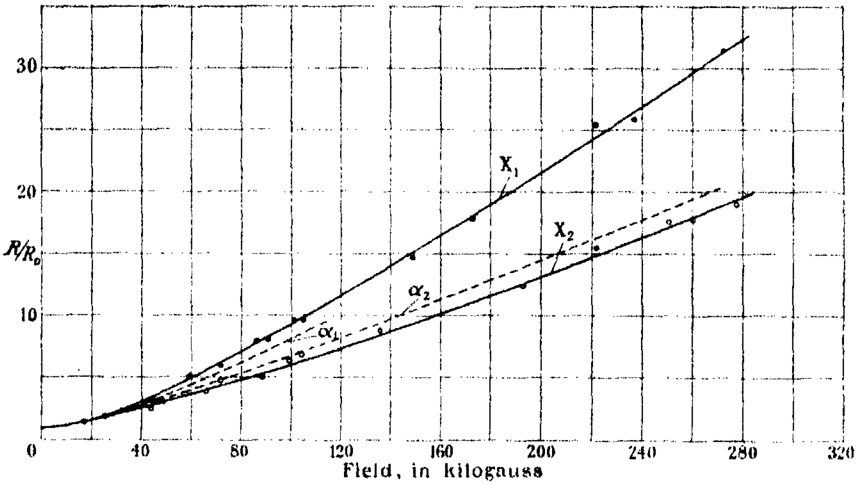


Fig. 14.—Room temperature, Orientation I $\perp$ Axis and I $\perp$ H. Curve X_1, Crystal X, Axis $\parallel$ H. Curve X_2, Crystal X, Axis $\perp$ H. Curve α_1 and α_2 for bismuth condensed from vapour in vacuum.

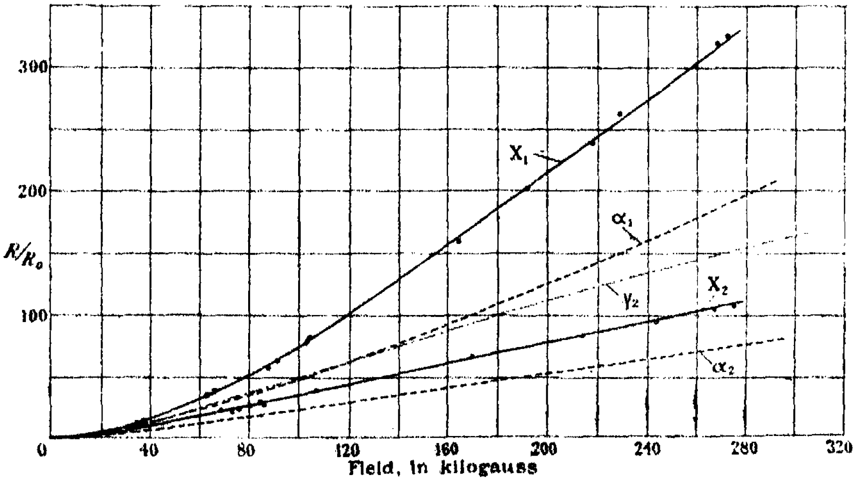


Fig. 15.—Temperature of the mixture of solid CO_2 and Ether, Orientation I $\perp$ Axis and I $\perp$ H. Curve X_1, Axis $\parallel$ H. Curve X_2, Axis $\perp$ H. Curve Y_2, Crystal Y, Axis $\perp$ H. Curve α_1 and α_2 for condensed bismuth.

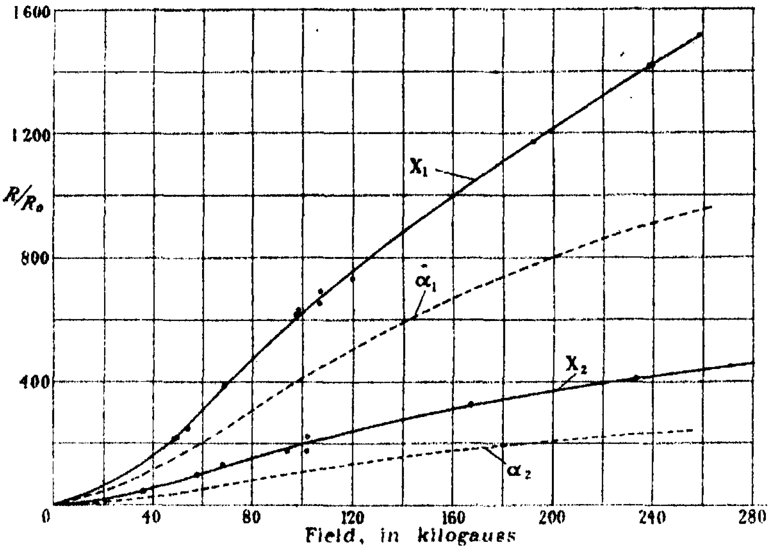


Fig. 16.—Temperature of Liquid Air, Orientation I ⊥ Axis and I ⊥ H. Curve X_1, Axis || H. Curve X_2, Axis ⊥ H. Curves α_1 and α_2 for condensed bismuth.

by curve X_1; in the second position (axis || H) the experimental results are represented on the same diagram by curve X_2.

The diagram of fig. 14 shows the results for the change of resistance at room temperature (15°). The crystal rod had the resistance of 0·00423 ohm. The change of resistance in this case is of the same magnitude and, in fact, is practically the same as curve A_2 on diagram fig. 11, where the current was perpendicular to the cleavage plane. Curve X_2 is for the same crystal rod when it is turned through an angle of 90° round its axis. The change of resistance is smaller, but the general character of both curves is the same. We have a curvature in the beginning which tends to a straight line at stronger fields.

Curves X_1 and X_2 in fig. 15 relate to the case where the crystal was placed in a mixture of solid CO_2 and ether. The resistance of the crystal was 0·00285 ohm. The curves are of an exactly similar character as at room temperature—they commence with a curvature and then straighten out. Finally, in fig. 16, we give the curve for liquid air temperature. Here the resistance of the crystal was 0·00155 ohm. We see that the general character of the curve is again the same. From a comparison of the results for the three temperatures we see that the difference between curve X_1, relating to the case where the cleavage plane was perpendicular to the magnetic field, and curve X_2, referring to the case where it was parallel, gets larger, as the temperature is lower. The magnitude of change of resistance for the position X_1 is of the same order as the one observed for the crystal A with the axis perpendicular to the cleavage plane (figs. 11, 12 and 13).

To see how the imperfection of the lattice affects the change of resistance for this orientation, we took a crystal which we had grown on a copper plate and which was probably far from being perfect, since its specific resistance was $1 \cdot 14 . 10^{-4}$. For this crystal, which we call Y, at room temperature we obtained the same change of resistance as given by the curves shown in fig. 14 (X_1 and X_2), but at the temperature of solid CO_2, where we made only one measurement in which the cleavage plane was parallel to the magnetic field, we obtained a larger change of resistance than for the crystal X, this being shown by curve Y_2 on fig. 15. This is probably due to the fact that in this case we had in the crystal Y, as suggested in Part I, local changes in the direction of the trigonal axis in the crystal lattice. This would account for both the large specific resistance and the departure of the curve Y_2 from X_2 since in the local imperfections, the cleavage plane will be more nearly perpendicular to the current and the resistance is thereby increased.

(3) *Discussion of the Results for the Experiments on the Change of Resistance when the Current is perpendicular to the Magnetic Field.*

From the results described it is easily seen that when the current is flowing perpendicular to the magnetic field there is *a great similarity in the change of resistance in the field for crystals differently orientated.* All the curves of figs. 11, 12, 13, 14, 15 and 16 are curved at the beginning, when the field is weak, and then at stronger fields straighten out. On closer examination we find that the initial part of each curve can be expressed by a quadratic law, viz.,

$$\Delta R/R_0 = \alpha H^2. \tag{1}$$

For the straight part of the curve we have

$$\Delta R/R_0 = \beta H. \tag{2}$$

It is convenient to describe each curve by the constant α and β, and also by the strength of the magnetic field at which the curve ceases to follow the square law (1) and changes to the linear law (2). This critical field strength, which we shall call H_K, can be given only very approximately, as in most curves no definite bends exist.

If we consider the influence of orientation of the crystal on these three constants, we see that for a definite *temperature the orientation of the crystal axis* *relative to the magnetic field chiefly affects the constant α of the square law, but the constant β of the linear law is the same for all orientations but one.* Thus the crystal A (fig. 12) at the temperature of solid CO_2 has the same slope of the linear part ($\beta = 1 \cdot 48 . 10^{-3}$) on both the curves, but in the orientation A_1 line $\perp$ H $\alpha = 1 \cdot 36 . 10^{-8}$ and in the orientation (line $\parallel$ H) curve A_2, $\alpha = 0 \cdot 8 . 10^{-8}$. At liquid air temperature the curves A_1 and A_2 (fig. 13) again have the same slope ($\beta = 5 \cdot 44 . 10^{-3}$). At room temperature (fig. 11) the slopes are different, but as is indicated by the slight curvature, it is very probable that at stronger fields (beyond 320 kilogauss) they will come to parallelism. In this case we have for the curve A_1, $\alpha = 1 \cdot 28 . 10^{-9}$ and $\beta = 1 \cdot 27 . 10^{-4}$, and for the curve A_2, $\alpha = 1 \cdot 4 . 10^{-9}$ and $\beta = 1 \cdot 53 . 10^{-4}$. For the crystal X with its cleavage plane along the axis of the rod, we have for the orientation (axis $\parallel$ H), curve X_1, exactly the same values for β as for the crystal A at the corresponding temperatures, but the α for the three temperatures under investigation has smaller values than for crystal A, viz., for room temperature $\alpha = 1 \cdot 24 . 10^{-9}$, for solid CO_2 $1 \cdot 0 . 10^{-8}$, and for liquid air $0 \cdot 88 . 10^{-7}$. The only exception is for the orientation (axis $\perp$ H) curve X_2. Here the part of the curve which obeys the square law is very short and the values for α and β are smaller than for the curves A_1, A_2 and X_1.

The effect of the temperature is primarily to produce a very rapid increase in α and β with the lowering of the temperature. From experiments at only three temperatures no definite law of change of α and β can be deduced, but we estimate from our measurements that α and β change inversely as the fourth or fifth power of the absolute temperature.

The effect of the chemical impurities of the crystal on the change of resistance may be described as only affecting H_K and β, making them smaller. From a comparison of the curves for the pure crystals made from Hartmann and Braun bismuth with the crystal B made of the less pure bismuth of Kahlbaum, we saw that the curves have the same initial change of resistance at all temperatures. This means that α is the same. But for the temperature of liquid air (*see* the corner of fig. 13) already at a field of $H_K = 8$ kilogauss, the curve B changes sharply to a straight line with a very small β ($1 \cdot 35 . 10^{-4}$). At the temperature of CO_2, this change point is not very well seen from fig. 12, as the scale of the diagram is too large, but actually it is given approximately by $H_K = 18$ kilogauss and the curve then follows a straight line with $\beta = 2 \cdot 8 . 10^{-4}$. At room temperature this change occurs only at a field of 260 kilogauss (fig. 11).

If now we consider *the influence of the perfection of the lattice, we see that it* *only affects the value of the critical field H_K*, the values of α and β being independent of the state of perfection of the lattice. This we see from a study of the curves for the crystals A, C, D and E, which were made from the same purest bismuth of Hartmann and Braun, but grown in different ways. It is seen from fig. 12 that at the temperature of solid CO_2, the temperature at which we were able to trace the influence of the imperfection of the lattice, all the different crystals having the same orientation have the same value for α within the limits of experimental error. This means that in the beginning they follow the same square law, but with the difference that the crystals which we take to be most perfect, such as the crystal E, follow this square law to higher fields, that is H_K is larger.

After the change to the linear law of increase of resistance, however, we have the same slope, or β is the same. For instance, the crystal E, at a temperature of solid CO_2, follows the square law fairly well to a field strength of 100 kilogauss, and the crystal which is less perfect only to 60 kilogauss. Similarly, we observe on the curves of fig. 13 taken at liquid air temperature, that the breakdown from the square law for the crystal E does not occur until 120 kilogauss, but for the crystal A_2 at a field strength of about 40 kilogauss.

To sum up all these results in a general way, we may think of the influence of the temperature, impurities and other factors, on the change of the resistance of bismuth in the magnetic field in the following way.

A crystal lattice can be distorted, first by the temperature, due to the vibration of the atoms, secondly by impurities, thirdly by the imperfections of the lattice, of the type described in the First Part of this paper, and finally by a strain set up in the crystal. At room temperature the disturbance of the lattice due to the thermal motion of the atoms plays a predominant part and the other factors have no marked influence. This explains why, for all our crystals at room temperature, we had similar curves to those given by A_1 and A_2 on fig. 11.

At the temperature of CO_2 the disturbances due to the impurities in the lattice, as in the case of the Kahlbaum bismuth crystal B, begin to be comparable with the disturbances due to the temperature motion of the separate atoms in the lattice, and we have the difference between pure and impure bismuth as seen from the curves on fig. 12. The lattice imperfections produced during the growth of the crystal are also comparable with the disturbance due to the temperature, and on fig. 12 we see the difference in the curves A, E, C and D.

Finally, at a temperature of liquid air the disturbances of the lattice, produced by the temperature motion, are so small in comparison with those due to the disturbances due to the imperfection of the lattice, that the latter factors are even more appreciable and may make a still greater difference in the curve for the change of the resistance, as seen on fig. 13. In this case, for example, it is even possible to trace the very small disturbances produced in the lattice by small strains set up through binding the holder with a silk thread, as was described in the last paragraph.

As the results obtained in our experiments depend upon the chemical impurities and upon the imperfections of the crystal lattice, it is important to discuss the question of how perfect our crystals were from a chemical and physical point of view, and how closely the change of resistance in a magnetic field, which, we observed, approaches that of an ideally perfect crystal. It is very difficult to answer these questions. As regards the chemical purity of the bismuth which we used, that of Hartmann and Braun is by reputation the purest, and it is possible that the purity was sufficiently high to make the results independent of the influence of chemical impurities. To settle this question systematic analysis and further purification of bismuth are necessary. It is unfortunately much more difficult to find out how perfect was the crystalline lattice of our best crystals. The difficulty lies in the fact that we possess at present only one independent means of judging the perfection of the crystal lattice, viz., X-ray analysis, but in our case it appears not to be sufficiently certain. We saw from the first part of this paper that the X-ray analysis shows us that a crystal such as E, which was grown in the uniform temperature condition (*see* fig. 5), is, as far as can be determined by the X-ray reflections, a perfect crystal, but the X-ray analysis fails to trace the existence of a few cracks which we know still exist in the crystal.

It is very probable that these cracks, if they are not very numerous, have no influence on the change of resistance in the case when the current is perpendicular to the perfect cleavage plane (I || axis). As, however, we are not sure of the number of cracks, the question of the perfection of the crystal lattice must be left open. The question is made more difficult by the fact that even if we admit that the crystal is nearly perfect after growth, it is very difficult to be sure that after cutting a piece 3 or 5 mm. long from the rod, and fixing it in the holder, the lattice is not disturbed. It was noticed that if the crystal was carelessly handled, *e.g.*, by over-heating when it is soldered between the silver discs *d*, fig. 9, or by pressing it too hard when the crystal is introduced in the ebonite holder *f*, fig. 9, the crystal was very often spoiled. This could sometimes be traced by an examination of the crystal under a microscope. The crystal appeared bent and had lines on the surface, which indicate that slipping occurred between the cleavage planes. We also noticed that such damaged crystals behaved, in a magnetic field, like crystals with an imperfect lattice. In the latest experiments we took the greatest care to handle the crystals carefully, but even then it was impossible to be sure that strains set up from the holder, due to the difference of thermal expansion between the ebonite and bismuth, do not spoil the lattice. It is very difficult to avoid all these disturbances produced in fixing the crystal, when such small pieces of the crystal are used. To avoid all these difficulties the experiments would have to be repeated with a coil of larger diameter, where a more satisfactory method of fixing the crystal would be possible. We think, however, that in cases where special attention was given to fixing the crystal (as in the case of crystal E), the distortion was not very great and the error introduced was of the same order as the general precision of our measurements.

In the following paragraph, where the change of the resistance of bismuth with the current parallel to the magnetic field is described, we shall again discuss the perfection of the lattice of our crystal, as in this case the imperfections of the lattice have an even larger influence on the change of resistance. It will be shown that it is even possible that all the effect is to be ascribed to the imperfections. If this be the case, these experiments afford another method of finding the perfection of the crystal.

Leaving aside the question of the absolute perfection of the crystals which we were using, we may still, from a comparison of the results from different crystals, obtain information as to how an ideally perfect bismuth crystal will change its resistance in a magnetic field.

First of all it is evident that even an ideal crystal in all orientations within the range of temperature which we investigated will first change its resistance with the square of the magnetic field and after the critical field linearly. Secondly, taking into account the influence of the impurities and imperfections of the lattice, it is evident that an ideal bismuth crystal in the region of weak fields, when the change of resistance follows the square law, will have the same coefficient α as our crystals in corresponding orientations and at corresponding temperatures. The difference between our crystal and a crystal made of ideally pure bismuth will be that the latter may have a larger value for the coefficient β for the fields where the increase of resistance follows the linear law, and a crystal which has a more perfect lattice than ours may have a larger value for the critical field H_K, which separates the region of the square and linear laws. This difference between an ideal crystal and our crystal may, however, be apparent only at lower temperatures, and we may be sure that for ordinary temperatures the two curves A_1 and A_2, fig. 11, and curves X_1 and X_2, fig. 14, will be practically identical with that of an ideal crystal, as the influence of the imperfection of the crystal is in this case small compared with the disturbances produced by the heat motions of the atoms.

For the experiments at a temperature of solid CO_2, we find that the impurities and the imperfections of the crystal lattice have already a marked influence, and all that we can say is that for our most perfect crystal E, H_K is about 100 kilogauss (whilst for the less perfect crystal A, $H_K = 60$ kilogauss), so that it is impossible from our data to fix the value of H_K for an ideal crystal, but it will certainly not be less than 100 kilogauss. The same argument applies to the linear part of the curve, the slope of which gives the coefficient β, and also to the experiments at the temperature of liquid air.

We shall now discuss one of the most interesting phenomena—the existence of what we have called the critical field H_K in the change of resistance of bismuth. This critical field is not well defined, but it may have an important theoretical significance. Below the critical field the crystal follows a square law of change of resistance, the coefficient α of which depends very much on the orientation of the crystal, but after the critical field we have a linear law which, for all orientations of the crystal [except one in the orientation for the curve (X_2)] is the same for any crystal made of the same bismuth and observed at the same temperature. This suggests that in the region of the linear change of resistance after the critical field, the condition for the motion of the conducting electrons is similar in all crystals for all orientations. It would appear, therefore, that above the critical field the influence of the neighbouring atoms on the conducting electrons, which depends on the symmetry of the crystalline lattice, is lost. This change or "reconstruction," as we may call it, produced by the field, whatever it may be, probably takes place during the increase of the magnetic field from O to H_K during the square law part of the curve. The ease with which the reconstruction takes place depends not only on the initial orientation of the crystal relative to the field, but also on the perfection of the crystal. We may say that impurities and imperfection in the crystal which disturb the lattice makes the reconstruction easier and reduces the value of H_K.

It will be interesting to see whether any other physical phenomena, such as diamagnetism, also change after the critical field is reached. If such a change be observed it will throw light on the character of the reconstruction, about which we can speak only vaguely at present.

As we may expect this reconstruction to be noticeable in the most favourable cases only at fields higher than 50 kilogauss, it is evident that it could not have been observed in the weak fields used by previous investigators.

Finally, we have to remark that it is difficult to compare the results just described with old ones. First, because in previous experiments the residual e.m.f. was never eliminated, and this, even in weak fields, as has been shown by Geipel,* reaches a value of 7 to 8 per cent. of the potential drop along the crystal. Secondly, because the crystals were not so perfect as those prepared by our methods, and we have seen how markedly the measurements are affected by the imperfection of the crystal lattice, and thirdly, because observations were confined to weak magnetic fields which were not covered in our investigations.

(4) *The Study of the Change of Resistance of Bismuth Crystals when the Current is parallel to the Magnetic Field.*

For the experiments described in this paragraph, we have been using a holder for the crystal given on fig. 9c. The results of the observation are given by curves in the same way as before, the scale for the ordinates being in this case much greater as the change of resistance is here very small. The letters by which the curves are marked refer to the same crystal as studied in the transverse phenomenon, but the suffixes in this case refer not to the orientation of the crystal but to the temperature. We use the suffix $_1$ for room temperature, $_2$ for solid CO_2, and $_3$ for the temperature of liquid air. We will first consider the set of observations relating to the curve when the current was perpendicular to the cleavage plane. We call this case I || H and I || axis.

In the first place we studied a piece of the same rod from which the crystal A was cut. The results are given on curves A_1, A_2 and A_3, fig. 17. Here we are at once struck by the fact that the resistance change is not only very small but quite definitely reaches a limiting value. For instance, at room temperature, curve A_1, fig. 17, shows that the resistance increases very little after reaching a field of 120 kilogauss. At a temperature of CO_2, the curve A_2 shows a slightly larger change of resistance, but the saturation phenomenon is even more marked and it is reached at a smaller field—about 80 kilogauss. The oscillogram, No. 10, taken for this curve, illustrates this saturation phenomenon. It can be seen very clearly that the deflection of the potential oscillograph reaches a limiting value and remains constant, although the current through the main coil is increased.

---

* 'Ann. d. Physik,' vol. 38, p. 149 (1912).

The most curious curve was obtained for liquid temperature (curve A_3). This shows not only this saturation phenomenon, but in addition the resistance,

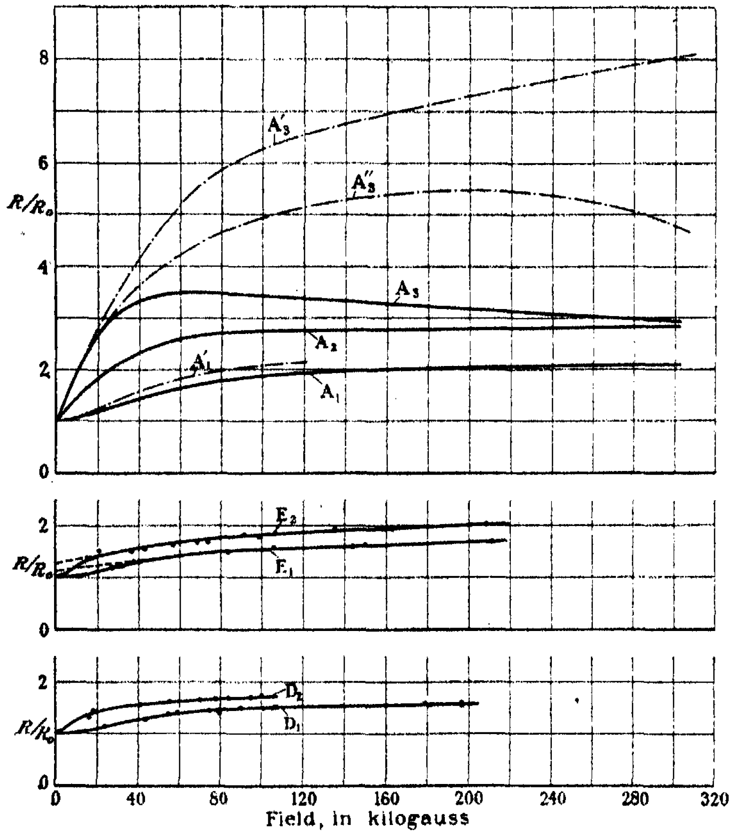


Fig. 17.—Orientation I ‖ Axis and I ‖ H. Curves A_1, A'_1, D_1, E_1 room temperature. Curves A_2, D_2, E_2, temperature of solid CO_2 and Ether. Curves A_3, A'_3, A''_3, temperature of Liquid Air.

after reaching a maximum, at about 60 kilogauss, gradually began to drop. The phenomenon was so curious that we decided to repeat the experiment. It happened that the first experiment was made towards the end of the week, when the liquid air used had been stored for a week. The experiment was repeated at the beginning of the next week with fresh liquid air. Unfortunately the contacts in the crystal were spoiled before this experiment, and to re-establish them the screw *g* on fig. 9 C was pressed down, thereby changing the stress conditions in the crystal. First of all we found the change of resistance in this experiment to be larger than before, as is seen from curve A'_3, but no drop of resistance was observed. Repeating the experiment with the old liquid air we obtained curve A''_3 and here again a drop of resistance is seen. The explanation of this phenomenon is very simple. When we used old liquid air most of the nitrogen had evaporated and we had a larger concentration of oxygen in the liquid air than in a fresh sample. It was mentioned in Part II of this paper that, owing to the larger paramagnetism of liquid oxygen, the pressure in the Dewar flask in very strong fields may be very considerable. For instance, if we had pure oxygen only, the pressure corresponding to a field of 300 kilogauss will be about 15 kilograms per cm.^{2}. With fresh liquid air this pressure will be only about 3 kilograms per cm.^{2}. Now it is evident that the fall in the resistance, curves A_3 and A_3'' can easily be accounted for by this pressure, if we assume that the cracks in the crystal close under the pressure produced by the liquid air, thereby diminishing the resistance. With the increased concentration of oxygen in old liquid air the pressure is larger and the closing of the cracks is more effective.

After the experiment with liquid air the curve was again taken for room temperature, and we obtained the curve A'_1. We see that the change of resistance is somewhat larger than for curve A_1 before the screw *g* was tightened up. When the crystal was taken out and examined under a microscope we found that the crystal rod was considerably bent, and the deformation of the crystal was quite evident. Apparently this deformation was produced when the screw *g* was screwed down.

For the next experiment we took a piece of the very flexible crystal and made a special holder; the leads (*e*), fig. 9 C, were arranged so that they were pressed to the silver discs from the sides, and practically no longitudinal pressure on the crystal was possible. With this flexible crystal we obtained curves D_1 and D_2. We see that the change of resistance is about one-half of that for the crystal A for the two corresponding temperatures. Using a piece of the rod from which the crystal E was made, grown, as stated, in the apparatus, fig. 5, and believed to be the best crystal, we again obtained a smaller change of resistance than for the crystal A at room temperature (curve E_1) and also at the temperature of CO_2 (curve E_2).

We were unable to perform the experiments for the crystals E and D at the temperature of liquid air for the reason that in the new holder, in which there was no axial pressure on the crystal, the soldering broke in the case of the crystal D when placed in liquid air, and in the case of crystal E the crystal itself broke. It has already been mentioned that it is very difficult to use these brittle crystals in liquid air, as most of them crack along the cleavage plane, if no pressure is applied to the crystal along its axis.

We shall now describe experiments with the other possible orientations of the crystal with its cleavage plane parallel to the axis of the rod (case I || H and H ⊥ axis). The crystal was placed in the same holder, fig. 9 C, and the results of the experiments are given by the curves in fig. 18. A piece of the same crystal rod was used as in the experiments with the current perpendicular to the field, and which we called crystal X. The curve marked by X_1 refers to room temperature, by X_2 to solid CO_2 and ether, and X_3 to the temperature of liquid air. The results from the point of view of magnitude and general character are exactly similar to those obtained in the case previously treated, except that the pressure effect of liquid air does not now appear. This we ought to expect as the possible cracks are parallel to the current, and their closing or opening will not influence the resistance of the crystal rod.

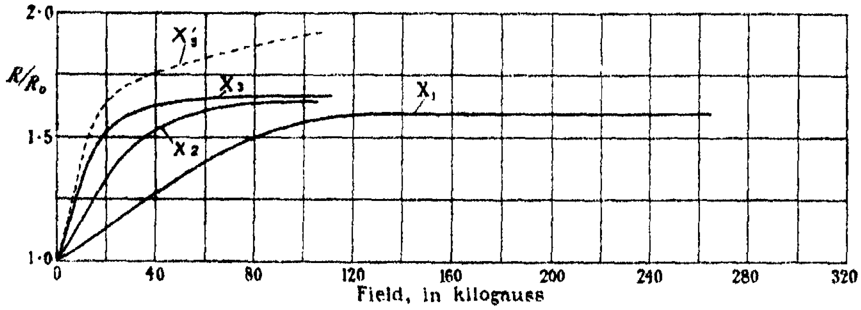


Fig. 18.—Orientation I ⊥, Axis and I || H. Curve X_1, room temperature. Curve X_2, temperature of solid CO_2 and Ether. Curve X_3, X'_3, temperature of Liquid Air.

From all these experiments we see that in this case we have an exactly opposite phenomenon to that observed in the experiments when the current was flowing perpendicular to the magnetic field. The change of resistance is much larger for the imperfect crystals than for the perfect ones, and in the case (I || axis) it is very easily influenced by the pressure. This all suggests a simple explanation of these phenomena. If we imagine that there is a crack, as in crystal A, or an imperfection, as in crystal X, which produces a large local increase in the resistance of the crystal, then the current in these places will not flow parallel to the axis of the rod, but will simply choose the way of the minimum resistance and may conceivably have local components which are perpendicular to the axis of the rod. Now when the crystal is placed in a magnetic field which is parallel to the axis of the rod, these perpendicular components of the current will flow perpendicular to the magnetic field, and the resistance, in the perpendicular direction, will be very much increased, as has been seen from the previous section. Finally, the current will be compelled to flow parallel to the rod, taking the paths which had a larger resistance, in the absence of the field. When this parallelism in the lines of flow of current is reached no further increase of the resistance of the bismuth rod is possible, and we have the saturation phenomenon as shown by the curves on figs. 17 and 18. If we have a crack in the crystal, which offers at a certain place a very large resistance, the current will flow round this crack, but when the crack is closed by means of pressure, as may happen in the case of liquid air, the current may flow through the broken part again, and it is evident that the resistance of the rod will diminish. It is highly probable that all the effect of the change of resistance, which we have observed in the case where the current is parallel to the magnetic field, is due to the imperfection of the crystal. This point of view, which reduces all the phenomena to a secondary effect, simply due to the imperfection of the crystals, is also confirmed by the fact that all the changes of resistance for the three different temperatures are of the same order. This fact alone makes it almost impossible to suppose that the phenomena have the same origin as in the case where the current is perpendicular to the magnetic field. On the other hand the observation that the saturation value is reached at a lower temperature with a smaller field, exactly supports our view, as at low temperatures, owing to the very rapid increase of the resistance for the small perpendicular components of the current with increasing magnetic field, the straightening out of the current parallel to the rod will happen more quickly. Our view is also strongly supported by the fact that at all the three temperatures the same relative change of resistance is observed. This is seen most clearly on fig. 18, where the curves X_1, X_2 and X_3 have the same saturation value. This is exactly what we should expect from our explanation. The path of the current without the magnetic field, as well as the changed path with the field, will be the same if the imperfections of the crystal do not alter for different temperatures, and this will evidently result in the same relative change of resistance.

If this explanation of the phenomenon is correct, we may use it for judging the perfection of the crystals. In our most perfect crystals, such as crystal E, we are still left with some increase of resistance. This is not necessarily due to the imperfection of the crystal, but may be produced by some imperfection of experimental arrangement. For example, at the point where the bismuth rod is soldered to the silver disc, the contact may not be equally good over the whole surface and this will result in the current being brought out of parallelism producing exactly the same result as the imperfection of the crystal. This was confirmed in some experiments in which after resoldering the crystal the limiting value for the change of resistance sometimes increased or diminished. The breaking of the contacts in liquid air, described in the previous paragraph, has also a very large effect on the increase of resistance.

In addition, part of the effect is no doubt due to the fact that it is impossible to cut out a piece of rod without damaging the crystal near the place of section. This effect is probably more marked in the crystal with the cleavage plane parallel to the rod, like crystal X, where more strain is put upon the crystal when it is cut, than in a crystal where the cutting is made along the cleavage plane.

The small rise in resistance after the saturation value is reached may be attributed to the fact that the crystal was not perfectly aligned with the magnetic field and the magnetic field was not perfectly uniform in that part of the coil where the crystal was placed. It is evident that if this is the case the current flowing in the crystal will tend to flow along the lines of force of the magnetic field, for if it flows perpendicular to the lines of force it will meet a very large resistance. With increasing field, the lines of flow of the current will gradually shift, and the resistance of the crystal will increase. On fig. 18 we have an illustration of this phenomenon. Crystal X was turned in the holder to make a small angle with the magnetic field, and instead of having the curve X_3 the curve X'_3 was obtained which has the same saturation value, but after which a steady increase of resistance is still well marked.

In our experiments, as we were working with small pieces of bismuth, a perfect alignment of the axis of the rod with the magnetic field was not obtainable. But we may roughly separate the increase of resistance due to this error in setting the crystal, from the other, if we extrapolate the flatter part of the curve as shown by dotted lines for the crystal E on fig. 17. Then the ordinate at which this curve intersects the axis of co-ordinates will give us the saturation value of the change of resistance which will occur in the same crystal when it is properly aligned with the field.

We see that the increase of resistance for the most perfect crystal E in this case will be only 15 per cent. at room temperature. At the temperature solid CO_2 and ether it will be about double. This is probably due to the strains set up in the crystal at lower temperatures owing to the difference of temperature expansion in the crystal which may open the cracks and affect the contacts between the silver disc and the bismuth rod. The change of resistance in crystal E is very small and it is quite possible that it may be fully accounted for by the imperfection of the experimental arrangement, as already described. It is very probable that by further refinement of the experimental arrangements we may be able still further to reduce the small change of resistance observed. It is important to mention that the change of resistance which we have observed is already considerably smaller than in previous observations.* A precise comparison is unfortunately difficult as these results all relate to weak fields (about 2 to 4 kilogauss) in the region where we can only extrapolate our data with an accuracy of 20 to 30 per cent. Even in this case, however, the values obtained in our experiment showed that the increase of resistance was several times (in no case less than three times) smaller than previously obtained, showing that our crystals, grown in the way described in the First Part of this paper, were much more perfect This is also confirmed by some other facts, for it is known that imperfect crystals give a smaller decrease of resistance when the temperature is diminished. In the table below we give the ratio of the resistance at the temperature of solid CO_2 (193° K.) and at the temperature of liquid air (91° K.) to that at room temperature (290° K.) for the two directions along and perpendicular to the crystal axis for our best crystals and compare them with the results for crystals which were grown in crucibles and studied by Borelius and Lindh in a magnetic field A.†

	I ∥ Axis.		I ⊥ Axis.	
	$\frac{R_{290}}{R_{193}}$	$\frac{R_{290}}{R_{91}}$	$\frac{R_{290}}{R_{193}}$	$\frac{R_{290}}{R_{91}}$
Our crystals	1·84	3·00	1·54	2·73
B. and L.	1·33	1·88	1·46	2·46

From this table it is evident that our crystals were more perfect than those obtained previously.

If our suggestion for the explanation of the phenomenon of change of resistance is correct, we see that our best crystal is crystal E, as we assumed in the previous section and as was confirmed by X-ray analysis.

---

* Borelius and Lindh, 'Ann. d. Physik,' vol. 53, p. 104 (1917); Schneider, 'Phys. Rev.,' vol. 31, p. 251 (1928).

† Borelius and Lindh, 'Ann. d. Physik,' vol. 51, p. 613 (1916).

(5) *The Change of the Resistance of Bismuth condensed from Vapour in Vacuum.*

As has been seen from the previous experiments at low temperatures, the influence of the imperfection of the crystal lattice plays an important part in the phenomenon of the change of resistance in magnetic fields.

Since we were unable to produce what we may call an ideally perfect crystal by the methods described in Part I, an attempt was made to produce perfect bismuth crystals by growing them by means of direct deposition of bismuth vapour on a plate. The general idea of this experiment was that if the bismuth atoms from the vapour were directly deposited on the crystal they would at once place themselves on a rhombohedral lattice, and all possible strains, due to the transformation from one modification to another, as described in Part I, would be avoided. These experiments were not successful, but some results obtained may be of interest.

The apparatus used in these experiments consisted of a quartz tube, 30 cm. long and about 1 cm. in diameter, having a bulb about 3 cm. from the bottom end, this end being sealed. Beneath the bulb the quartz tube was drawn to a neck of 3 or 4 mm. inside cross section. During the experiment the tube was evacuated to the highest possible vacuum by means of charcoal and liquid air. Some pure Hartmann and Braun bismuth was placed in the bottom of the tube. About 1 cm. above the neck, in the middle of the bulb, a quartz plate, 8 mm. in diameter, was placed on a solid copper holder. The holder was fixed to a glass tube inserted in the quartz tube from the open end in such a way that the disc with the holder did not touch the quartz tube. When the bismuth in the bottom of the tube was heated a molecular beam was formed which struck the quartz plate and formed a deposit. Arrangements were provided to warm up the copper holder by means of an electrical heater, or to cool it with water or liquid air. A thermo-couple, touching the quartz disc, recorded the temperature.

When the deposit was obtained the following curious phenomenon was observed. If the temperature of the quartz plate was low—at the temperature of liquid air, or even at room temperature—the bismuth deposit had no apparent crystalline structure. If the bismuth deposit was ½ mm. thick, it formed a grey mass which was very brittle and in no way resembled ordinary bismuth. The specific resistance of this bismuth was very high, and the temperature coefficient was negative.

It appears from these experiments that at low temperatures the atoms stick to the plate without forming a crystal and most probably form an amorphous mass. We noticed that on the surface of the quartz bulb, which kept fairly warm during the experiment, the layer deposited was of a different character, and was much more like ordinary bismuth. This suggested heating the quartz plate to a higher temperature by means of the electric heater. In this way we were really able to obtain layers which were of a quite definite crystalline character. When we kept the plate at 200° C. the bismuth deposit obtained no longer had a negative temperature coefficient, and its specific resistance was about $1 \cdot 7 \,.\, 10^{-4}$. Depositing bismuth throughout the day at such a temperature we obtained a deposit of about 3 to 4 mm. thick which looked very like a little mountain made of needles stuck perpendicular to the quartz plate. These needles slightly adhered to one another, and it could be seen under the microscope how these needles broke away when pressed by the point of a sewing needle. Each individual needle apparently consisted of other needles, which, by the slight pressure of a pin, could be separated. It was clear that the structure of the deposit was crystalline, consisting of many very small needles having only one general orientation in the direction of the incident beam of molecules.

The following attempts were made to obtain single crystals instead of a bunch of needles. In the first place, instead of the quartz plate, a large bismuth crystal was used with its freshly broken cleavage plane perpendicular to the beam, but the deposit grew on it in an exactly similar way as on the quartz plate. This might be due to impurities absorbed on the surface layer of the bismuth crystal. The following experiment was then made to obtain a clean surface. After a small layer of bismuth had been deposited on a quartz plate it was heated above the melting point of bismuth. The surface layer then broke up into small drops which remained stuck to the plate, and when cooled each drop turned to a single crystal providing a centre of crystallisation for the condensing bismuth atoms. As the whole procedure took place in a high vacuum, the possibility of having a contaminated surface layer was less likely. The bismuth deposit was then made on the drops, as before, at a temperature of 200°, but on examination it was seen that the cleavage plane of these drops had no relevance to the small needles which stuck to them, and that there were about 50 or 60 needles growing from one drop. This means that the orientation effect on the crystalline axis, produced by the direction of the falling molecules, is much more important in the growth of the deposit than the plane of the crystal on which it is deposited.

However, a plate about 1 mm. thick and 3 mm. wide was cut out from the

3 mm. thick deposit made of needles, and was placed in the holder in order to investigate its change of resistance in a magnetic field. The results obtained are shown by curves α_1 and α_2 in fig. 14 for room temperature, in fig. 15 for a temperature of solid CO_2, and in fig. 16 for the temperature of liquid air. These curves all relate to the case when the current is perpendicular to the magnetic field. In the case described by curve α_1 the current passed through the plate under examination, parallel to the direction of the needles. In the case of α_2 the current passed perpendicularly to the needles. The same plate was used for obtaining both these curves, only it was soldered in the holder in different ways. We see that the change of resistance in this case, especially at low temperatures, is smaller than for the bismuth crystal in our experiments. The reason for this is probably that the boundary layers between the individual needles, which are probably quite good crystals, play the rôle of imperfections in the crystal lattice. This point of view is confirmed by the fact that curve α_1 is always above curve α_2, because in the former case the current is passing along the needles, so that it has fewer boundaries to cross.

The phenomenon here observed bears a close relation to the observation made by Patterson* and afterwards studied by Hollivequet† and studied in more detail by Hecker and Curtis.‡ They found that a very thin bismuth layer, deposited by cathodic sputtering on a cooled plate, does not change its resistance in a magnetic field. The deposited layer commences to show a change of resistance in magnetic fields only after it has been heated to a higher temperature. This phenomenon is apparently not connected with the thinness of the layer, but is simply due to the fact that bismuth deposited by cathodic sputtering on a cold surface probably forms an amorphous mass in the same way as by vapour deposition, which may be described as a crystal spoiled to its limit, and in this amorphous state bismuth has no change of resistance in a magnetic field. As soon as it is heated, the layer crystallises itself and the normal bismuth properties appear. This is confirmed by some of our observations in which much smaller changes of resistance in magnetic fields were found for layers about ½ mm. thick, when these layers were deposited on a cooled plate.

(6) *The Residual E.M.F. occurring in Bismuth in the Magnetic Field.*

The residual e.m.f. developed in bismuth in a magnetic field, when a current is passed, reached a very considerable value in our experiments as is easily seen from oscillograms 2 and 3. Our experimental method was not suited to a study of this phenomenon with great accuracy, since, as is seen from the oscillogram, the inverse e.m.f. does not reach its limiting value in 1/100 second. We therefore made only a general study of the magnitude of this phenomenon in different field strengths, for different orientations of the crystal and for varying temperatures. The magnitude of the residual e.m.f. was estimated in the following way. We used the field with the flat top, as shown on oscillograms Nos. 2 and 3, Plate 5, and broke the current in the crystal in a quite definite place towards the end of the flat part of the current wave in the coil by means of the special switch described in Part II. Then from the oscillograms we measured the ratio of the negative deflection of the potential oscillograph, which is due to the residual e.m.f., to the full amplitude, at the moment of break of the current through the crystal. This ratio, which we call n, is evidently given by

---

* 'Phil. Mag.,' vol. 64, p. 562 (1902).

† 'Phil. Mag.,' vol. 64, p. 135 (1902).

‡ 'Phys. Rev.,' vol. 15, pp. 65 and 457 (1920).

$$n = \frac{\mathrm{E}}{\mathrm{IR}}, \tag{3}$$

where E is the residual e.m.f., R the resistance of the crystal in a magnetic field, and I the current through the crystal. From the value of n and R we were able to judge the absolute magnitude of E. It was at once noticed that this e.m.f. occurs only in the cases where the current was flowing perpendicularly to the magnetic field. The maximum e.m.f. for crystals is observed in an orientation when the current is perpendicular to the perfect cleavage plane, the phenomenon being several times smaller, when the current is parallel to the cleavage plane. This result is in agreement with the previous observations made by Geipel.* In this research we limited our study to the case where the current is perpendicular to the cleavage plane, as only in this case could we expect a reasonable accuracy from our method of measurement of the e.m.f.

In the first instance we verified that n is independent of the current sent through the crystal. This was easily done by taking an oscillogram, as shown on oscillogram 2, then doubling the current through the crystal and taking another oscillogram with the same magnetic field in the coil, but halving the sensitivity of the potential oscillograph. When the two oscillograms were superposed the complete coincidence of the lines traced by the potential oscillograph proved that the ratio n does not depend on the current density, and that E is proportional to I. We also proved that n does not depend on the length of the crystal rod. Using a holder for the crystal, shown in fig. 9, A, we first measured n for a given field between the potential leads and afterwards between the silver discs, in this way doubling the length of the crystal. The value of n obtained for these two experiments was the same. These results are in agreement with those of Seidler[*] and Geipel,[†] for smaller current densities and for weaker fields.

---

* 'Ann. d. Physik,' vol. 51, p. 503 (1916).

The first result which we will consider is the change in the e.m.f. produced by changing the orientation of the crystal when it is turned round its axis by 30°. We took a crystal in the two orientations described in para. 2 (line $\parallel$ H and line $\perp$ H), and it was easily noticed that the ratio n is larger in the case when the crystal is placed relative to the magnetic field in such a way that one of the lines on the perfect cleavage plane was perpendicular to the lines of magnetic force (line $\perp$ H), for which the crystal has also its maximum change of resistance. For instance, for one of the crystals, which we call K, at room temperature, at a field of about 190 kilogauss, in the orientation (line $\perp$ H), the ratio $n = 0\cdot18$, and for the same magnetic field, but after being turned by 30° to the position (line $\parallel$ H), the ratio $n = 0\cdot10$. These two cases are illustrated in oscillograms 11 and 12, both taken under similar conditions, oscillogram 11 being for the position (line $\perp$ H) and oscillogram 12 for (line $\parallel$ H). It can easily be seen that n is greater in oscillogram 11. This influence of the orientation for a given crystal remains unaltered for all strengths of the magnetic field and for different temperatures.

The influence of temperature and the field strength on the e.m.f. depends very much on the perfection of the crystal, the chemical and crystalline perfection of the crystal influencing the e.m.f. even at room temperature. We observed that the most perfect crystals, which at low temperatures give the largest increase of resistance, also have the largest values for the ratio n. We shall first consider the influence of the field and temperature on the value of n for one of the best crystals K. The value of n at each temperature was measured for two or three field strengths, and we found that at room temperature, for the orientation (line $\perp$ H), and a field $H = 95$ kg., $n = 0\cdot175$; for $H = 191$ kg., $n = 0\cdot18$. For the orientation (line $\parallel$ H), $H = 95$ kg., $n = 0\cdot09$, and for $H = 194$ kg., $n = 0\cdot10$. At the temperature of solid CO_2 for the orientation (line $\perp$ H) we had for $H = 97$ kg., $n = 0\cdot22$, and for $H = 194$ kg., $n = 0\cdot19$ (oscillogram 2). At liquid air temperature different specimens of a good crystal gave the following values for n: at $H = 95$ kg., $n = 0\cdot06$; $H = 156$ kg., $n = 0\cdot075$, and $H = 247$ kg., $n = 0\cdot04$. From these data we see first of all that n increases slightly at the temperature of solid CO_2 and drops to a small value at liquid air temperatures. For each of these three temperatures the value of n does not seem to depend much on the field strength. The small drop in n, which is observed with increasing strength of the field, may be accounted for by the fact that in strong fields the time to develop the full e.m.f. may be longer. The experimental error is also considerable in these experiments as the negative deflection due to the e.m.f. is small. In most cases the estimated error in the absolute value of n is $\pm 0{\cdot}01$.

---

[*] *Loc. cit.*

[†] *Loc. cit.*

From the expression (3) for n, we see that if n is constant, E will be proportional to R. This means that the residual e.m.f. will change with the magnetic field in the same way as the resistance, as shown in the curves, figs. 11, 12 and 13. Thus in weak fields the e.m.f. will be proportional to the square of the magnetic field and will afterwards increase linearly with the field strength.

The imperfections of the lattice are found to exert a great effect on the absolute value of n, but as in the case just described, n changes but little with the magnitude of the magnetic field, even for a relatively imperfect crystal. If we compare the oscillogram No. 9, Plate 7, which was taken for one of our best crystals E, with the oscillogram No. 7, Plate 6, taken at the same temperature, but for a less perfect crystal, we see that the negative deflection of the potential oscillograph is nearly twice as large for the crystal E as for A. This shows that the most perfect crystal from the point of imperfections of the lattice showed the largest e.m.f. For instance, at room temperature, when the crystal was placed in the maximum position (line $\perp$ H), different specimens of bismuth crystal, grown in different ways from the same Hartmann and Braun bismuth, gave values of n ranging between $n = 0{\cdot}13$ and $n = 0{\cdot}19$ (for the crystal A, $n = 0{\cdot}14$); at a temperature of solid CO_2 and ether n ranged from $0{\cdot}22$ to $0{\cdot}14$ (for the crystal A, $n = 0{\cdot}15$).

The chemical impurities prove to have an enormous effect on the e.m.f. The less pure crystal B, already described in para. 2, which was made from Kahlbaum bismuth, gave at room temperature the following values for n. For the maximum position (line $\perp$ H), where $H = 95$ kg., $n = 0{\cdot}14$; where $H = 194$ kg., $n = 0{\cdot}17$ (oscillogram No. 13, Plate 8); in the position (line $\parallel$ H) where $H = 93$ kg., $n = 0{\cdot}07$, and where $H = 195$ kg., $n = 0{\cdot}10$, these values being only slightly on the small side. For the temperature of solid CO_2, however, we have the following data. In the maximum position B (line $\perp$ H), where $H = 195$ kg., $n = 0$—no noticeable e.m.f. In the minimum position B_2 (line $\parallel$ H) where $H = 95$ kg., $n = 0\cdot03$ (oscillogram No. 14, Plate 8), and where $H = 195$, $n = 0$—again no e.m.f. This suggests, if we examine the curves B_1 and B_2 on figs. 11 and 12, that as long as the change of resistance of the impure crystal B follows the curves of the pure crystals A, C and D, the e.m.f. has a more or less normal value, but as soon as the field reaches such values that the curve of B no longer follows the curve A, the e.m.f. not only ceases to increase, but gradually drops to zero.

As previously stated, the e.m.f. has already been studied by Seidler and Geipel,* and a theoretical explanation of this effect has been given by Heurlinger.† Their experiment was carried out by means of a ballistic galvanometer which was switched on to the bismuth wire after the main current had been broken. If we compare their results with ours, we find that the maximum value of n which they observed, if expressed in our way, was $n = 0\cdot08$, which is about three times smaller than our maximum value. This is probably due to the greater imperfections in the lattice of their crystals. Our results for the change of the e.m.f. with the temperature, when expressed in the same way as those of Seidler or Geipel, do not quite agree with theirs. These investigators state that the maximum e.m.f. is observed about the temperature of solid CO_2 and ether. In our case this was only true for imperfect crystals, the best crystals showing a continuous increase in e.m.f. with decreasing temperature. For instance, if we take the crystal K, taking the e.m.f. to be unity at room temperature, in a field of 95 kilogauss at the temperature of solid CO_2 $E = 13$, and for the temperature of liquid air $E = 50$. The crystals with a bad lattice do not show any appreciable e.m.f. at the temperature of liquid air, and this is probably why Geipel observed his maxima for the e.m.f.

We very roughly studied the decrease of the e.m.f. with the time from oscillograms like oscillogram 3, where the current was switched off at the beginning of the flat part of the current wave. The value obtained was of the same order as observed by Seidler and Geipel, and as would be expected from the theory.

The explanation of the origin of the residual e.m.f. as given by Heurlinger‡ can be described in the following way. It is known that if a current is passing through a conductor in a magnetic field a difference of temperature occurs across the conductor in a manner very similar to the Hall effect. This phenomenon was discovered by Ettinghausen and is called after his name. As a result of this transverse temperature gradient, a thermal current is set up perpendicular to the bismuth rod. On the other hand, Nernst has discovered that perpendicular to a thermal current in a magnetic field a difference of potential will occur. In our experiments with bismuth we have the two phenomena occurring at once. The difference of temperature produced perpendicular to the rod, due to the Ettinghausen phenomenon, results in a thermal current which produces the longitudinal e.m.f.

---

* *Loc. cit.*

† 'Phys. Z.,' vol. 17, p. 221 (1916).

‡ *Loc. cit.*

As it appears from the study of the Ettinghausen phenomenon that it takes some time to establish the temperature gradient, this accounts for the time lag phenomenon. From the data given by Zahn[*] for the Nernst and Ettinghausen phenomenon, Heurlinger obtains the correct value and sign for the resulting e.m.f. in bismuth. A verification of this theory is impossible in our case as the magnitude of the Nernst and Ettinghausen phenomenon is not known for the field strengths with which we are working. As, however, the general theory of the phenomenon is no doubt the same, we find from this theory a full justification for the elimination of the e.m.f. in the way stated in section 1 of this paper. It is evident that this e.m.f. is only a secondary phenomenon, which unfortunately reaches very large values for good crystals in strong magnetic fields. (For instance, in the case of measurements made in liquid air in a field of about 155 kilogauss, when a current of $\frac{1}{2}$ ampere was sent through the crystal, the e.m.f. reached the value of 0·16 volt in the length of 3 mm.)

As already stated, our experiments were of an approximate character only, but as the residual e.m.f. is only a secondary phenomenon, depending on the separate effects of Ettinghausen and Nernst, it was not thought worth while to refine our experiments. We hope later to make a separate study of these two effects in strong fields, but even from our present experiments it is quite evident that the Ettinghausen and Nernst effects in bismuth depend not only on the orientation of the crystal but also markedly on the state of perfection of the crystal. Further, from the constancy of n for a definite crystal in a variable magnetic field which makes the change of resistance proportional to the residual e.m.f., we must also admit a close physical connection between the phenomena of Nernst and Ettinghausen and the change of resistance of bismuth in magnetic fields.

If the difference of temperature due to the Ettinghausen phenomenon, at a given temperature, in a crystal of given orientation, is

$$\Delta T = f(H)\, I \qquad (4)$$

[*] See note *, p. 436.

where I is the current and f (H) a function depending in this case only on the field H, then in the same bismuth crystal, at the same temperature, the longitudinal e.m.f. due to the Nernst effect will be equal to

$$E = \Delta T \,.\, F(H) \tag{5}$$

where F is another function characterising the Nernst phenomenon. We get from (3), (4) and (5) $F(H) f(H)/R = n =$ constant for a given crystal at given temperature.

From measurements of Zahn* for weak fields we know that $f(H) \propto H$ and $F(H) \propto H$. Thus we get $R \propto H^2$ as we actually observe in weaker magnetic fields. Since for stronger fields $R \propto H$ we must have

$$F(H) \,.\, f(H) \propto H.$$

This shows that in strong fields the Ettinghausen and Nernst phenomena cannot both be proportional to the magnetic field as is usually assumed, and as is true for weak fields.

(7) *Comparison with Theory.*

The theory of the change of resistance of conductors in a magnetic field was first given by J. J. Thomson,† then by Ganz‡ and recently by Sommerfeld,§ but it is very difficult to account by means of any of these theories for the large change of resistance which takes place in such a peculiar substance as bismuth.

In Sommerfeld's theories of electrical conductivity it is assumed that free electrons exist in the metal, but, as a result of the strong interaction between the electrons, they have not the classical distribution of velocities, but a different one given by the modern quantum and wave mechanical theory. The law of distribution of velocity derived from the quantum theory, as given by the Fermi-Dirac statistics, makes the velocity of the electrons much higher than on the classical theory, and also shows that the velocity distribution in a metal under ordinary conditions is practically independent of the temperature. This is the main difference between the new and the older theories. The picture of a current passing through a metal remains the same, namely, that when a difference of potential is applied to a conductor the free electrons get an excess of momenta in the electric field which is absorbed by the nuclei of the heavy

---

[*] 'Jahr. d. Rad.,' vol. 5, p. 166 (1908).

[†] 'Rapports, Congrès internat. Phys.,' vol. 3, p. 138 (1900).

[‡] 'Ann. d. Physik,' vol. 20, p. 293 (1906).

[§] 'Z. d. Physik,' vol. 47, p. 43 (1928).

atoms which form the lattice of the crystal. The specific resistance of the conductor depends on the efficiency of the mechanism of transmission of the excess of momenta from the electrons to the lattice. The effect of the magnetic field on a conductor after the theory of Ganz and Sommerfeld, consists in the fact that the electrons, due to the deflection of their paths by the magnetic field, accumulate less momenta between two collisions from the applied electric field, and the result is an increase of resistance of the conductor. The magnitude of the change of resistance in a magnetic field, calculated on the basis of this picture, is very small, and even for a substance like silver, which changes its resistance very little in a magnetic field, is, as was pointed out by Sommerfeld,* about ten thousand times less than the experimentally observed value. For bismuth, the discrepancy between the theory and experiments is even larger. If, for instance, from the formula 77A, given by Sommerfeld, we approximately calculate the free path of an electron necessary to account for the change of resistance in magnetic fields, taking the values for the coefficient of change of resistance (1), at liquid air temperature of $\alpha = 2 \cdot 76 . 10^{-7}$ (Sommerfeld's B), we find that the value l of the free path is above 1 cm., which seems impossibly large.

Further discrepancies between the theory and experiment lie in the fact that the theory gives the change of resistance to be proportional to the square of the field. For bismuth this is only true for weaker fields (1), but for higher fields we have the linear relation (2). We shall shortly publish some results for the change in resistance of other substances, and here again we find that the change of resistance in strong fields follows more closely the linear law. For instance, copper at liquid air temperature, in a field of 300 kilogauss, increases its resistance by approximately 35 per cent., and this increase follows very closely the linear law.

This difference, not only in magnitude of the effect, but also in the way in which it depends on the field, shows the difficulty of accounting for the change of resistance of a conductor in magnetic fields by deviation of electrons moving freely between two collisions. A more likely picture for the change of resistance in magnetic fields can be obtained if we assume that the magnetic field has a direct influence on the mechanism of collision between the electron and the atoms. It is quite evident that the picture given at the beginning of the paragraph, for the mechanism which accounts for the resistance of a conductor, is only a very crude one and is usually adopted only as an illustration. We see from the value of the free path of an electron of a conductor, as obtained from Sommerfeld's theory* that an electron passes a few hundred atoms before making a collision in which all the excess of energy and momentum is given up to the atom. Probably what actually happens is that the electron is continually transferring small amounts of its excess energy and momentum to the atoms, and the effect of all these small collisions will, at a distance which is called the free path, be equivalent to a complete collision.

---

* *Loc. cit.*, p. 56.

Our suggestion is that the disturbance produced by the magnetic field in the atoms of the conductor will increase the efficiency of these " small collisions," and this will increase the resistance. This hypothesis will not affect all the important advantages of Sommerfeld's theory, and, from a formal point of view, the effect of a magnetic field may be regarded as a decrease in the length of the free path of the electron.

The elaboration of this suggestion is difficult, as the mechanism of a collision between an electron and an atom is at present unknown in detail, but from the attempts already made† we may expect‡ to have this question solved on the basis of the wave mechanics. At present all that we may attempt to do is to compare the collision between an electron and an atom in a conductor with that which we observe in gases. Here we have the well-known Ramsauer phenomenon. This effect, so far as it has been studied, has two definite factors. First, that the interchange of energy, between an electron and an atom may be much smaller than may be expected from a collision theory; and secondly, that the efficiency of the collision depends to a great extent upon the structure of the atom. The most symmetrical atoms, like those of the inert gases, absorb very little momentum and energy from the electron in a collision, and for them we have the most complete Ramsauer effect.

To develop our suggestion further, we shall also assume that the free electrons in the metal make the most efficient collisions with the atoms of the crystalline lattice when they have an unsymmetrical structure.

By means of this hypothesis we may obtain an explanation of several phenomena observed in conductors. First of all we can explain why the specific resistance of bismuth is high and is so sensitive to the magnetic field. From our hypothesis we should expect that those elements, which have the most symmetrical configuration of the core when the free electron is removed, will make the least effective collisions with the electrons and will have the highest conductivity per atom. In the first column of the periodic system we may expect to have the most symmetrical atomic-cores, as they are similar to the inert gases, and we actually observe the highest conductivity per atom.[*] In bismuth the state of affairs is very different. It has been suggested by Ehrenfest,[†] from a consideration of the diamagnetism of bismuth, that the electrons move round several nuclei forming large orbits. These orbits will make the crystal lattice behave as though it were made up of molecules having a very unsymmetrical structure, which will be very unlike the structure of inert gases, and we have to expect a very small "Ramsauer effect," and this will account for the high specific resistance of bismuth.

---

* *Loc. cit.*, p. 24.

† L. Mensing, 'Z. f. Physik,' vol. 45, p. 603 (1927).

‡ Ehrenfest and Rutyers, 'Naturwiss.,' vol. 11, p. 184 (1928).

From our point of view we regard the effect of the magnetic field in increasing the resistance of a metal, as originating from its disturbing effect on an atom in changing the velocity of the electrons and orientating the orbits, making it thereby less symmetrical, and thus increasing its efficiency of collisions with electrons. Owing to the large size of the electronic orbits in bismuth the magnetic field will be much more efficient in spoiling the symmetry of the molecules, and in this way may very easily account for the exceptionally large increase of resistance. The experimentally observed influence of the orientation of the crystal lattice relative to the magnetic field on the change of resistance in bismuth may be explained by the fact that these large electronic orbits must have a definite orientation in the crystal, as the diamagnetic susceptibility of bismuth is different in different directions relative to the crystal axis. They are therefore differently affected by the magnetic field when it is applied at a different angle, and the efficiency of collision is changed leading to a different increase in the resistance.

The existence of a critical field (section 3) which separates the region of weak fields, where the change of the resistance is quadratic from the fields where it is linear may probably be explained by the fact that above H_K the critical value for the orientation effect of the magnetic field on the orbits is stronger than that of the inter-molecular forces arising from the symmetrical distribution of the atoms in the lattice, and the orbits reach a steady orientation. In this state the orbits are probably less affected by the field and have a smaller (linear) change of resistance. This is also strongly supported by the observed fact that the change of resistance after the critical field is reached in most cases does not depend on the initial orientation of the crystal.

The temperature effect may be described as follows. The atoms in a metal conductor, or, in the particular case of bismuth, the molecules, suffer a thermal oscillation in the lattice and disturb each other. This disturbance is evidently equivalent to making the atom less symmetrical and this increases the efficiency of the collision with the free electrons, the disturbances evidently being larger at higher temperatures. This will account for the decrease of resistance at low temperatures. As the amplitude of vibration is directly connected with the absolute temperature, we may expect that the decrease of resistance will also be a function of the absolute temperature. The combined effect of the magnetic field and the temperature is more difficult to picture, but it is conceivable that the two phenomena produce disturbances of a different character in the atoms which are not additive. At room temperature in bismuth the disturbances due to the temperature motion of the atoms prevail, and we have a smaller influence of the magnetic field. The lower the temperature the more prominent are the disturbances produced by the field.

---

[*] *See* Bridgman, 'Reports Solway Conference,' p. 70 (1927).

[†] *Loc. cit.*, Part I.

Impurities and imperfections of the lattice also have a similar effect on the conductivity. A foreign atom introduced in the lattice will disturb the symmetry of the lattice and each atom will be less symmetrically surrounded by others and the symmetry of the atom itself will be diminished. Since one foreign atom may disturb quite a large portion of the lattice, we can see why the value of the specific resistance is so sensitive to traces of certain impurities. A genuine imperfection of the lattice will act in a similar way. It is evident that at high temperatures the asymmetry produced by this imperfection will be small compared with that produced by the temperature motion of the atoms and will have less effect on the specific resistance and the change of resistance in magnetic fields (*see* section 2).

We have seen that all the conductivity phenomena of bismuth are more sensitive to impurities and imperfections of the lattice than those of any other metal. This can be explained by the fact that any imperfection not only affects the symmetry of the molecules, but may also break the large orbits into small ones. These have different properties and are less affected by the magnetic field, as is seen from experiments with the amorphous deposits. Here the large Ehrenfest orbits probably no longer exist, and no marked change of resistance in magnetic fields occurs. (It would be interesting to test the magnetic properties of the amorphous bismuth; on our assumptions it must be much less diamagnetic than in crystals.)

In sketching out these suggestions, we have always spoken about electrons in collision with atoms. We took this picture as it is more familiar. No doubt, however, any quantitative development of these ideas, if possible, can only be attempted on the basis of wave mechanics. In this case we have to replace the word electron by the word wave, and write, instead of the "efficiency of collision," the "scattering power of an atom," and so on. The chief assumption, however, upon which all this speculation depends will be in the wave mechanics that the scattering power of an atom in a metal is increased by the disturbance produced in the atom by a magnetic or any other kind of field, as, for example, that of neighbouring atoms.

Summary of Part III.

A method is described by which it is possible to eliminate the time lag phenomenon in the measurement of the change of resistance of bismuth in a magnetic field.

Experimental results are given for the change of the resistance of bismuth crystals when the current is perpendicular to the magnetic field, for different orientations of the axis of the bismuth crystal relative to the magnetic field.

The measurements have been made at 290°, 193° and 91° absolute.

It has been shown that the impurities and imperfections in the crystal lattice greatly influence the change of the resistance of bismuth. This is specially noticeable at low temperatures and in strong magnetic fields.

The change of resistance in a magnetic field of an ideal perfect crystal is discussed. It has been shown that in weak fields the change of resistance follows a square law and in strong fields it follows a linear law. The linear law of the change of resistance is practically independent of the orientation of the crystal relative to the magnetic field.

The case when the current is parallel to the lines of force of the magnetic field was also investigated for different orientations of the crystal axis relative to the magnetic field and a "saturation effect" was found. The change of resistance in this case is very small, is not much affected by the temperature, and its magnitude is strongly affected by the imperfection of the crystal. It is suggested that this small change of resistance is due to the imperfection of the crystals and to the defective aligning of the current in the crystal with the direction of the lines of the magnetic field.

Experiments are described in which it was attempted to make perfect bismuth crystals by direct deposition of bismuth vapour in vacuum on cooled plates. The experiments for obtaining single crystals were unsuccessful, but by studying the change of resistance in a strong magnetic field of the deposits obtained, it was possible to clear up certain phenomena previously observed in thin bismuth layers.

A rough investigation was made of the residual e.m.f. which occurs in bismuth in a magnetic field when a current is passing and which accounts for the time lag phenomenon. It is shown that the ratio of the residual e.m.f. to the resistance is independent of the strength of the magnetic field in the region of investigation, but is strongly affected by the orientation and by the perfection of the crystal.

A general discussion is given on the failure of the present theory of metallic conductivity to account for the phenomenon of the change of resistance in a magnetic field and in this connection some suggestions are made.

All the work produced in the three parts of this paper has been done with the continuous assistance of Mr. E. Laurmann, to whom I would like to express my personal thanks.

I am indebted to Mr. J. D. Cockcroft for the correction of the MS.

My thanks are also due to Prof. Sir E. Rutherford for the kind interest he has shown during the process of the work.

The work was all carried out in the Magnetic Laboratory with the support of the Department of Scientific and Industrial Research.

EXPLANATION OF PLATES 5 TO 8.

(Description of the Oscillograms.)

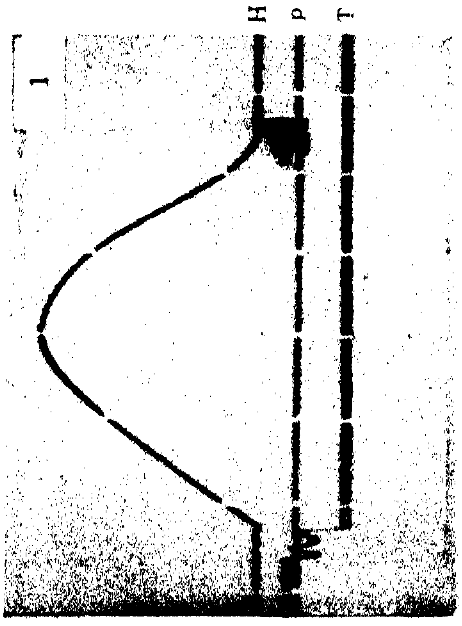
In all the oscillograms the time scale is from the right to the left. The interruption in the lines seen on some of the oscillograms are caused by interruption of the beam of light falling on a mirror of the three oscillographs by a little shutter operated by a synchronous motor. The interval of time between two interruptions is 1/180 of a second. One of the objects of the interruption is to mark corresponding points on the different curves.

Curve H is traced by the oscillograph which records the current through the coil where the magnetic field is obtained. Curve I is traced by the oscillograph which records the current through the crystal. Curve p is traced by the oscillograph connected to the potential leads of the crystal. In some experiments it was found necessary to measure the tension across the break of the main current through the coil, this curve is marked by T.

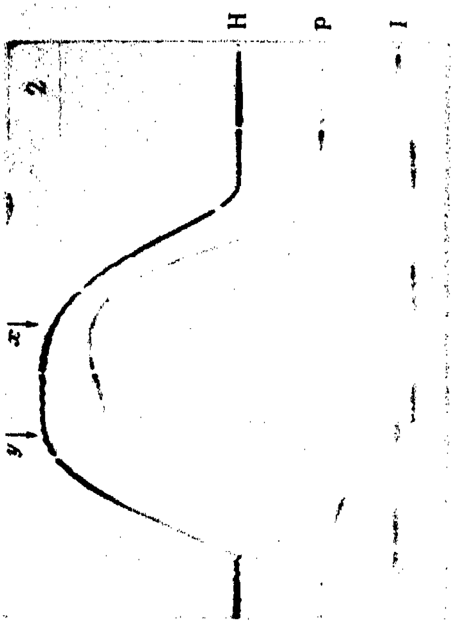
OSCILLOGRAM No. 1.—This oscillogram was taken to illustrate the induction in the circuit of the potential oscillograph due to the variation of the magnetic field. For this purpose no current was sent through the crystal and the traces of a deflection of the potential oscillograph were due to the induction. Holder fig. 9 B with the crystal B was used. Temperature 22·5° C. The field on the top of the current wave $H_{max.} = 320$ kilogauss.

OSCILLOGRAM No. 2.—($I \perp H$) and ($I \parallel$ Axis), crystal K. Holder fig. 9 B. (Line $\perp$ H). Resistance of the crystal $R_0 = 0{\cdot}00331\ \Omega$. Temperature $T = 193°$ K. (solid CO_2 and ether). $H_{max.} = 194$ kilogauss. Current through the crystal $I = 1$ amp. Resistance w (fig. 7) of the potential oscillograph $r = 20{\cdot}5\ \Omega$. Ratio between the residual e.m.f. and the potential drop due to the change of resistance $n = 0{\cdot}19$.

1

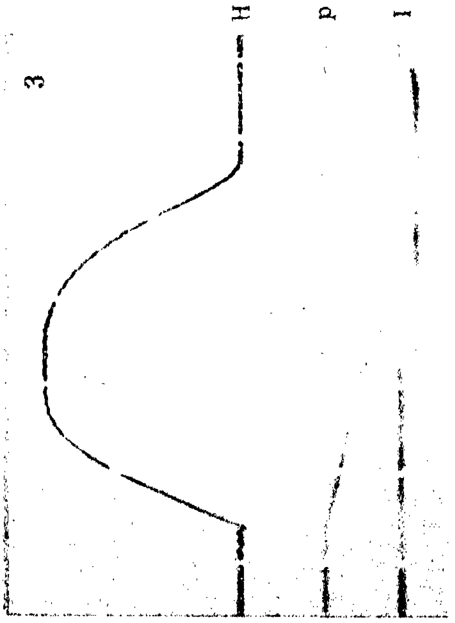


2

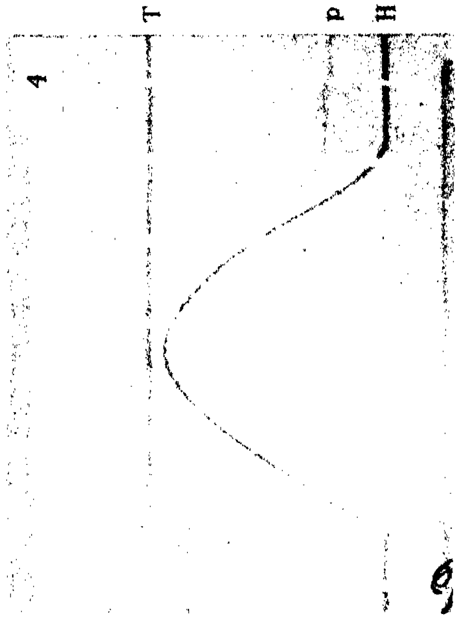


← Time

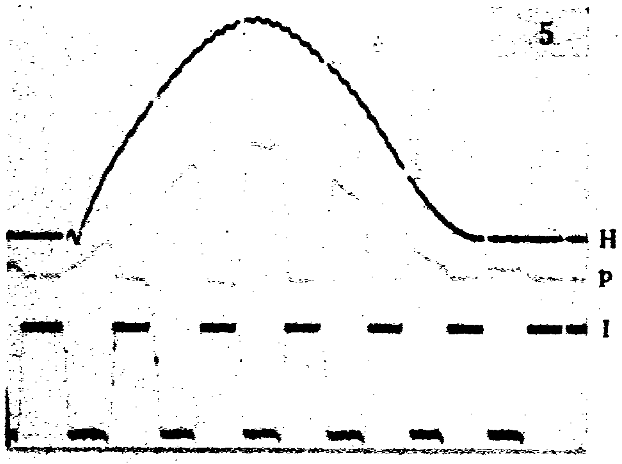
3



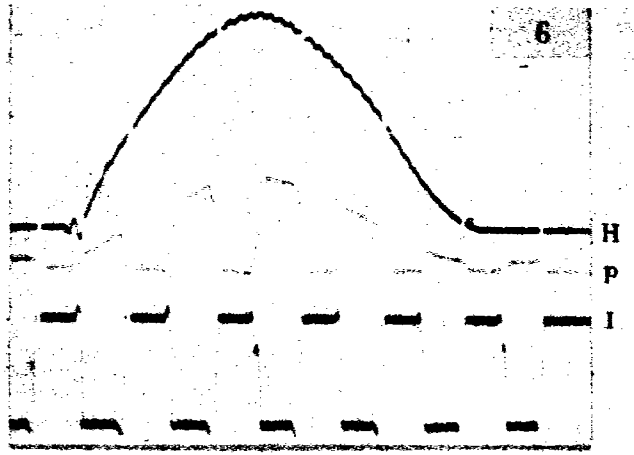
4



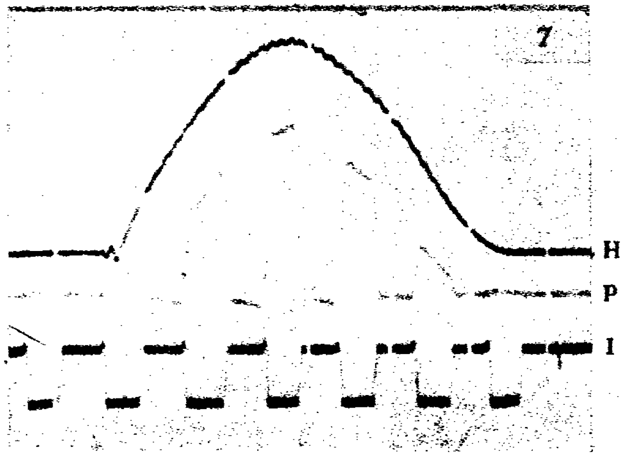
5



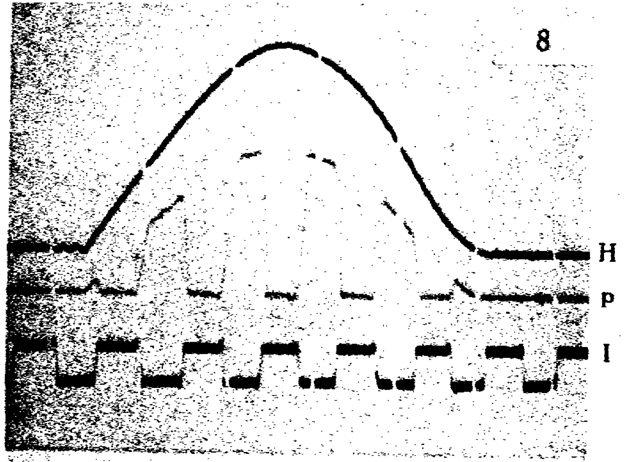
6



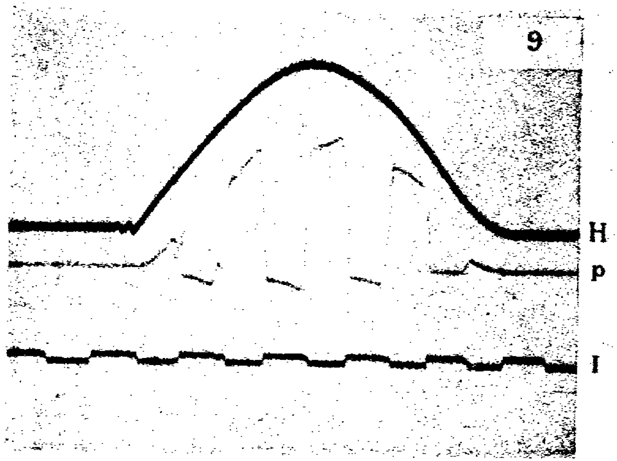
7



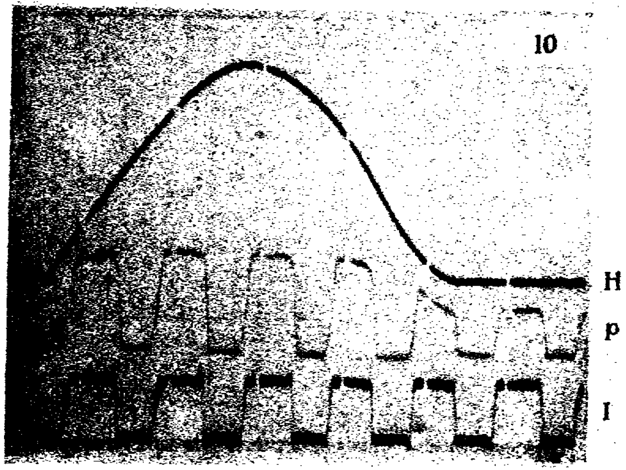
8



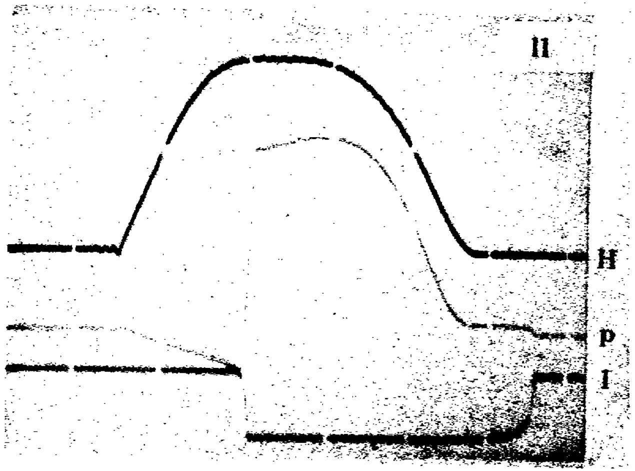
9

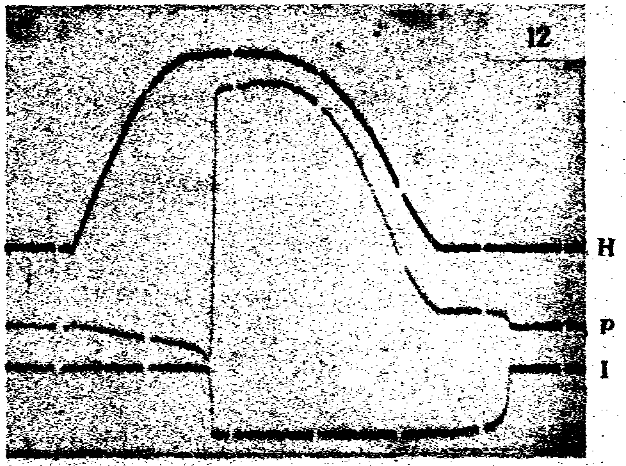


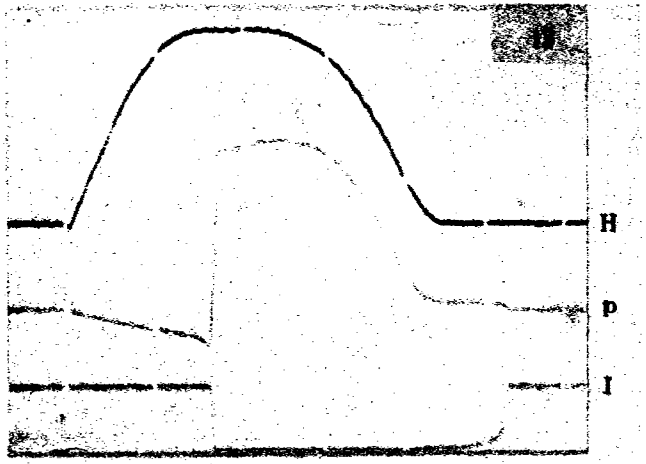
10



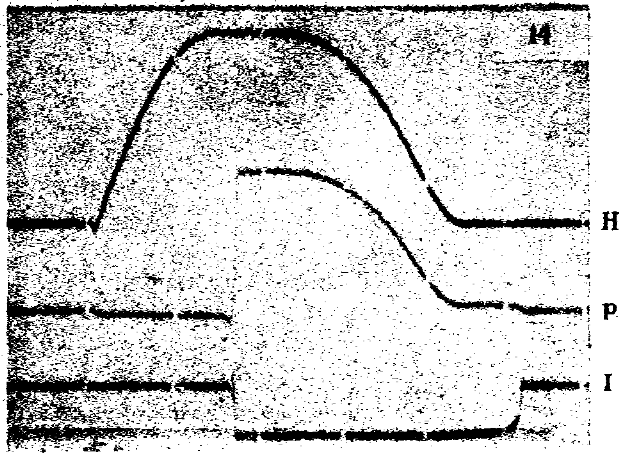
11




12


13

← Time →


14

Oscillogram No. 3.—The same conditions as oscillogram No. 2, with the exception that in this case the current through the crystal is broken earlier, $n = 0 \cdot 16$.

Oscillogram No. 4.—From this oscillogram the loop curve on fig. 10 is taken. $(I \perp H)$ and $(I \parallel \text{Axis})$. The crystal was cut from a bismuth plate grown like the rods in the apparatus on fig. 1. Holder fig. 9 A. $R_0 = 0 \cdot 0075\ \Omega$. Temperature 22° C. $H_{max.} = 242$ kilogauss. $I = 3$ amp.

Oscillogram No. 5.—$(I \perp H)$ and $(I \parallel \text{Axis})$. Curve A_1, fig. 11 (Line $\perp$ H). Holder fig. 9 B. $R_0 = 0 \cdot 00325\ \Omega$. Temperature = 17° C. $H_{max.} = 109 \cdot 5$ kilogauss. $I = 3$ amp. $r = 5 \cdot 05\ \Omega$.

Oscillogram No. 6.—All conditions similar to those in oscillogram No. 5 with the exception that the crystal is turned by 30°. Curve A_2, fig. 11. (Line $\parallel$ H), $H_{max.} = 108$ kilogauss.

Oscillogram No. 7.—Crystal A in the same position as on oscillogram No. 5 $(I \perp H)$ and $(I \parallel \text{Axis})$. Curve A_1, fig. 12 (Line $\perp$ H). Holder 9 B. $R_0 = 0 \cdot 00205\ \Omega$. Temperature = 193° absolute. $H_{max.} = 107$ kilogauss. $I = 1 \cdot 5$ amp. $r = 10 \cdot 05\ \Omega$.

Oscillogram No. 8.—Crystal A in the same position as in No. 5 $(I \perp H)$ and $(I \parallel \text{Axis})$ Curve A_1, fig. 13. (Line $\perp$ H), holder fig. 9 B. $R_0 = 0 \cdot 00112\ \Omega$. Temperature = 91° absolute (liquid air). $H_{max.} = 300$ kilogauss. $I = 1 \cdot 0$ amp. $r = 60 \cdot 05\ \Omega$.

Oscillogram No. 9.—Crystal E. $(I \perp H)$ and $(I \parallel \text{Axis})$. Curve E_1, fig. 12. (Line $\perp$ H). Holder fig. 9 B. $R_0 = 0 \cdot 00805\ \Omega$. Temperature = 193° absolute. $H_{max.} = 280$ kilogauss. $I = 0 \cdot 5$ amp. $r = 60 \cdot 32\ \Omega$.

Oscillogram No. 10.—Crystal A. $(I \perp H)$ and $(I \parallel \text{Axis})$. Curve A_1, fig. 17. Holder fig. 9 C. $R_0 = 0 \cdot 001055\ \Omega$. Temperature = 22·5° C. $H_{max.} = 300$ kilogauss. $I = 3$ amps. $r = 4 \cdot 05\ \Omega$.

Oscillogram No. 11.—Crystal K. $(I \perp H)$ and $(I \parallel \text{Axis})$, the crystal in this case being turned by 30° (Line $\perp$ H). $R_0 = 0 \cdot 0053\ \Omega$. Temperature = 23·5° C. $H_{max.} = 191$ kilogauss. $I = 3$ amps. $r = 10 \cdot 05\ \Omega$. $n = 0 \cdot 183$.

Oscillogram No. 12.—Crystal K. $(I \perp H)$ and $(I \parallel \text{Axis})$. (Line $\parallel$ H). Holder fig. 9 B. $R_0 = 0 \cdot 0053\ \Omega$. Temperature = 23·5° C. $H_{max.} = 194$ kilogauss. $I = 3$ amps. $r = 5 \cdot 05\ \Omega$. $n = 0 \cdot 10$.

Oscillogram No. 13.—Crystal B. $(I \perp H)$ and $(I \parallel \text{Axis})$. Holder fig. 9 B. (Line $\perp$ H). $R_0 = 0 \cdot 00463\ \Omega$. Temperature = 22° C. $H_{max.} = 194$ kilogauss. $I = 3$ amps. $r = 10 \cdot 05\ \Omega$. $n = 0 \cdot 17$.

Oscillogram No. 14.—Crystal B. $(I \perp H)$ and $(I \parallel \text{Axis})$. Holder fig. 9 B. (Line $\parallel$ H). $R_0 = 0 \cdot 00385\ \Omega$. Temperature = 193° absolute. $H_{max.} = 93$ kilogauss. $I = 3$ amps. $r = 10 \cdot 05\ \Omega$. $n = 0 \cdot 03$.

The Polarisation of Radio Waves.

By J. HOLLINGWORTH, M.A., M.I.E.E. (of the National Physical Laboratory).

(Communicated, by permission of the Radio Research Board, by Sir Joseph Petavel, F.R.S.—Received February 25, 1928.)

The following paper is primarily an attempt to study by experimental methods the problem of the propagation of long radio waves between two points on the surface of the earth. In the early days of radio-telegraphy the propagation formulæ in general use were admittedly empirical; being derived from the simple expressions for propagation over a plane conducting surface, and modified by factors deduced from the statistical mean of large numbers of observations. While such formulæ served their immediate practical purpose with considerable success, it was evident from their origin that they could give but little physical insight into the actual process of propagation. Further experiments were carried out with a view to increasing the amount of data available; but their general result was to indicate that the process was far too complex to be expressed in any reasonably simple mathematical form, and the investigation then gradually assumed the more theoretical form to which attention is being directed at present.

Now it was suggested in the early days that it was extremely probable that some form of electrical layer in the upper atmosphere was responsible for many of the phenomena; and in view of the more recent work by later observers, especially Smith-Rose and Appleton, it may be taken that this view has been definitely confirmed; it is indeed now tacitly accepted by the great majority of investigators. Starting from this point, mathematical investigations have been developed, including those of Eccles,* Larmor,† and Nichols and Schelling;‡ most of which take into account in addition the effect of the earth's magnetic field. But in the formulæ obtained in this way appear certain constants of the upper atmosphere to which at present no definite numerical values can be assigned, owing to our limited knowledge of the electrical and physical conditions at heights greater than 8 to 10 kilometres.

Owing to the presence of these unknowns, the most important of which is the variation of ionic density with height, these formulæ cannot be used for the direct calculation of received radio intensities ; in fact the most promising line of attack seems to be in the other direction ; namely to use the experimentally measured values as a means of checking the validity of the general assumptions, and also, if possible, for the determination of some of the abovementioned constants. It is with this end in view that the experiments here described have been planned and carried out, and while they only touch the fringe of the subject, which becomes more and more complex the further it is investigated, they seem to have yielded a few definite facts which must be taken into account in any theory of propagation.

---

* 'Roy. Soc. Proc.,' A, vol. 87, p. 79 (1912).

† 'Phil. Mag.,' vol. 48, p. 1025 (1924).

‡ 'Bell System Tech. J.,' April, 1925.

Method.

The absolute measurement of received radio intensities is a subject which has probably made less progress than most other radio measurements, owing to the technical difficulties involved. In the first case the received powers are extremely small, of the order of 10^{-12} watt ; so that considerable amplification is essential, with its attendant problems of instability and parasitic disturbances.

Secondly, the actual production for purposes of calibration of extremely small high frequency e.m.fs. of accurately known value has not yet been placed on a standardised basis ; almost each experimenter using a different method, the methods themselves as yet having been only mutually compared to a small extent. The earliest observations of the intensity of received radio signals, consisted almost entirely of simple measurements of the electric field intensity from very distant transmitting stations, usually at least 2000 km. away. This, of course, gave immediate practical information in its most useful form ; but such results were not always of the highest theoretical value, for which the experiments should be designed to produce the physical facts under observation in their most marked and easily measurable form.

It was therefore necessary for the present purpose to make very considerable modifications in apparatus and routine, though both of these have been matters of gradual growth rather than abrupt transition from the older ways. Early experiments carried out by the author on the old lines, though made at considerably shorter distances (300 to 1000 km.) had shown conclusively that several independent variables were involved, and that therefore adequate information could not possibly be obtained from single simple intensity measurements. In this connection it is important to point out that work of this nature differs from most other physical work in that the independent variables, even if their existence is fully understood, are incapable of control and independent variation, since they are really of the nature of cosmic phenomena. Moreover, on long wave-lengths at any rate, it is impossible to vary transmitter conditions as regards wave-length, situation, and time of sending, to the extent that pure theory demands; the only method is to make the best use possible of the standard commercial transmissions. Progress is therefore slow and laborious. The origin of the experiments in this paper dates back to January, 1924, when four simple intensity measuring sets were installed under the auspices of the Radio Research Board of the Department of Scientific and Industrial Research at Slough, Manchester, Aberdeen, and Glasgow (later transferred to Exeter), to work to a routine supplied to them by the author. The analysis of the results received from these stations over a period of about two years provided further evidence in favour of an interference effect between a surface wave and one returned from the upper layer, and this was subsequently confirmed by a special experiment; but during the winter period (November to April), marked, though systematic irregularities occurred, which were clearly inexplicable on the basis of the simple assumptions which had so far been found sufficient. About this time, papers by Eckersley[*] and in particular Smith-Rose and Barfield,[†] suggested that the errors in direction-finding on distant stations, which also become more pronounced annually during the winter period, were due to abnormal polarisation of the downcoming wave. Efforts to collate these errors with the above intensity variations led, however, to no definite results; and it was Naismith's discovery in May, 1926, that the daily cycle of intensity variation, which occurred regularly on certain stations during the hours around sunset, was invariably accompanied by an equally distinct and definite cycle of bearing variation, which first established the connection in a measurable form. Experiments were immediately put in hand to try and put this on a quantitative basis, and preliminary results were communicated in a letter to 'Nature.'[‡] The results of these experiments, analysed on lines shortly to be described, are given in fig. 1, but their numerical accuracy is obviously low, although comparison with later results (fig. 4) shows that their main features are correct.

At this stage it was felt that further qualitative evidence was unnecessary, and efforts were henceforward concentrated on the improvement of experimental method and accuracy.

---

[*] 'Radio Review,' vol. 2, p. 60 (1921).

[†] 'Roy. Soc. Proc.,' A, vol. 116, p. 682 (1927).

[‡] 'Nature,' vol. 118, p. 409 (1926).

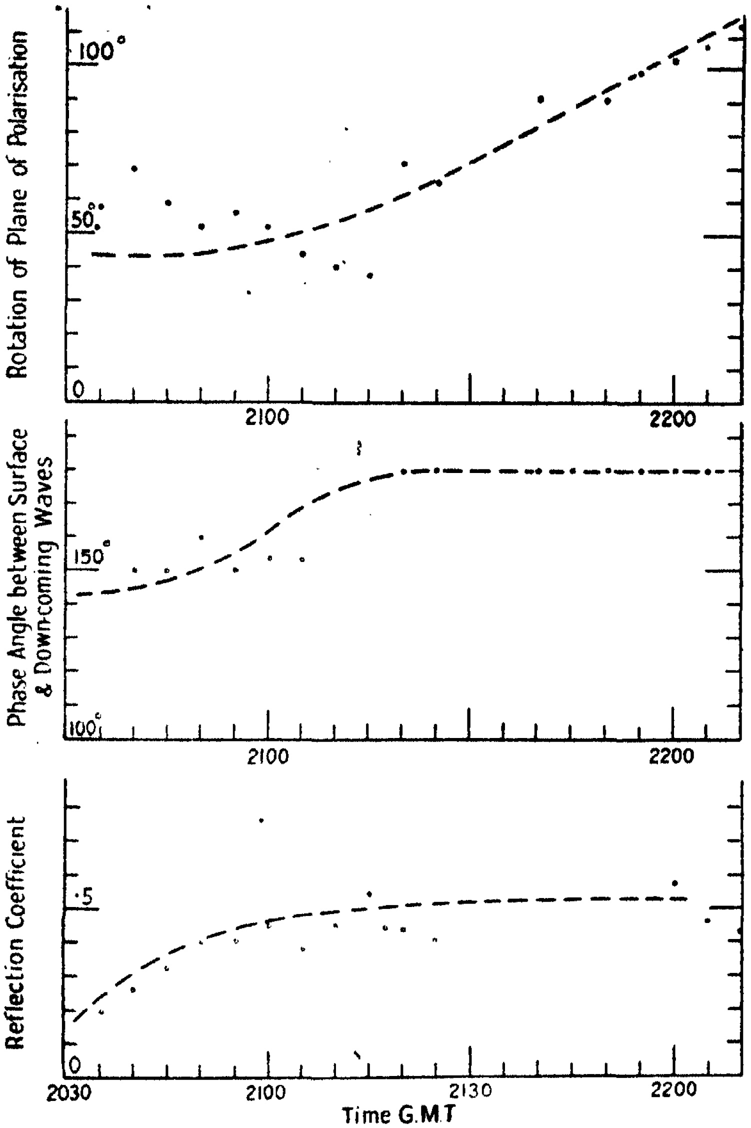


FIG. 1.—Analysis of Sunset Run on St. Assise (14,350 m.). Sunset 2009 G.M.T. (This was the first ever obtained and was taken with the old apparatus on June 28, 1926.)

Theory of the Method.

Since this work is not concerned at the moment with the actual physical mechanism by which the effects are produced at the upper layer, it will be sufficient to start from the general assumption that the down-coming wave is abnormally polarised, *i.e.*, that it contains a horizontal component of electric force; and that it is the interference between this complex wave and the surface wave which produces the effects under consideration.

The necessary equations will therefore contain six unknowns :—

(I) H_0 the magnetic intensity of the surface wave.

(II) H_1 the magnetic intensity of the normal component of the downcoming wave.

(III) H_1^1 the component of magnetic intensity in the vertical plane containing the direction of propagation.

(IV) θ_1 the phase angle between H_0 and H_1.

(V) θ_1^1 the phase angle between H_0 and H_1^1.

(VI) The "equivalent height" of the layer.

H_0, H_1 and H_1^1 are all assumed to be of simple harmonic form in time. Of these H_0 can be calculated from one of the standard diffraction formulæ, and (VI), the height of the layer is roughly known from Appleton's* work and from previous work of the author. It is also assumed, following the work of Smith-Rose,† that at these wave-lengths the surface of the earth may be assumed to be a perfect reflector.

Four unknowns being left, it will therefore in general be necessary to take four distinct measurements for each determination.

Now with this wave system the resultant horizontal magnetic intensity at right angles to the plane of propagation is

$$H_0 + 2H_1 \text{ (vector addition)},$$

and the resultant horizontal intensity in the plane of propagation is $2H_1^1 \sin \theta$ where θ is the angle which the direction of propagation of the downcoming wave makes with the surface of the earth. Hence the e.m.f. induced in a vertical coil oriented at an angle ϕ to the geographical direction of the arriving signal will be proportional to :—

$$(H_0 + 2H_1) \cos \phi + 2H_1^1 \sin \theta \sin \phi \text{ (vector addition)}.$$

Examination of this equation will show that by variation of ϕ, the only parameter under experimental control, three, and only three, separate equations can be obtained. Appleton‡ obtained a fourth on his shorter waves by the artifice of wave length variation; but since the percentage change required is proportional to the wave length, this is impossible in the present case on practical grounds. In theory this fourth equation can also be obtained by the use of a vertical aerial in conjunction with the coil ; but unfortunately this method is only applicable when the transmitter is at a comparatively short distance (<200 km.) from the receiver, since the new equation is the same as the one of the others except for a factor $\cos\theta$ in one term, which, if θ is small, only differs from unity by an amount comparable with the instrumental accuracy of the whole apparatus.

---

* 'Nature,' vol. 115, p. 333 (1925).

† 'Roy. Soc. Proc.,' A, vol. 107, p. 587 (1925).

‡ 'Roy. Soc. Proc.,' A, vol. 117, p. 576 (1928).

Therefore to obtain a determinate solution from three observations it is clearly necessary to make a further assumption. The one taken is that the down-coming wave is not elliptically but plane polarised, though with its plane of polarisation rotated from the normal position. (This must be taken to include the possibility of elliptic polarisation, the eccentricity of which cannot be detected owing to the limits of instrumental accuracy.) While on purely general grounds this assumption is quite unjustifiable, the results so far obtained yield a considerable amount of indirect and circumstantial evidence in its favour. In order, however, to develop the arguments on this point, it is first necessary to describe in some detail the method of obtaining the solution and the particular form in which the three readings are now taken.

On the assumption of plane polarisation if H^1 be the magnetic intensity of the down-coming wave, η the angle of rotation of its plane of polarisation from the normal position, the equations for the e.m.fs. induced in two vertical coils set 45° on either side of the geographical direction are

$$\left.\begin{aligned}\sqrt{2}E_1k &= H_0 + 2H^1\cos\eta + 2H^1\sin\eta\sin\theta\\ \sqrt{2}E_2k &= H_0 + 2H^1\cos\eta - 2H^1\sin\eta\sin\theta\end{aligned}\right\}$$

(vector addition) where k is an instrumental constant.

If therefore we can construct a triangle (fig. 2) OAB such that

$$OA = \sqrt{2}E_1k \quad\text{and}\quad OB = \sqrt{2}E_2k$$

and join O to C the mid point of AB it follows that

$$OC = H_0 + 2H_1\cos\eta$$

$$CA = CB = 2H_1\sin\eta\sin\theta.$$

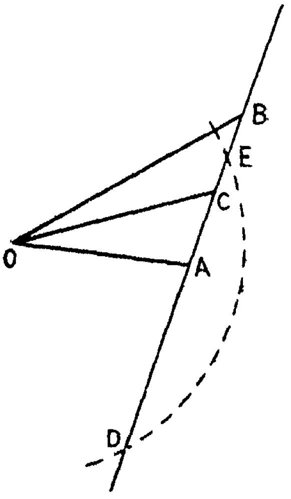


FIG. 2.

Hence AB gives the phase position of H_1 and therefore by describing a circle with centre O and radius H_0 to cut AB in D and E either CD or CE gives $2H_1 \cos \eta$ in magnitude and phase. (In any particular case the choice between CE or CD can be settled on inspection by a method which is somewhat elaborate to explain and not essential to the present argument.)

It is therefore evident that in every case the solution is obtained by the construction of the triangle OAB.

Now following out Smith-Rose's idea, since the primary symptoms of abnormal polarisation are the absence of a sharp minimum on a direction-finding coil and an error in the apparent bearing, the first three observations selected were :—

(i) The intensity in the normal (maximum) position of the coil.

(ii) The intensity in a direction at right angles to this.

(iii) The bearing error.

These being taken as rapidly after one another as possible. The results in fig. 1 were actually obtained from such observations; but it was soon found that their accuracy was low. In the first place the intensity (ii) is often very much smaller than intensity (i), involving either a loss of valuable time while the apparatus is readjusted, or else a reading of low accuracy for observation (ii) affected severely as it is by jamming owing to its low intensity. Also, when abnormal polarisation is large, the minimum becomes very flat, resulting in an extremely low accuracy of observation (iii). This method was therefore replaced by one in which three intensity observations at different orientations were made. Again, however, the accuracy was low, partly due to the fact that the apparatus then in use was not sufficiently refined for the purpose, but also owing to a purely geometrical factor. For in this case we have to construct the triangle OAB given OA OB and OC; and when these are nearly equal and AB is small compared with them (as often happens in practice) the shape of the triangle is excessively sensitive to small changes in these lengths. It was, therefore, finally decided to construct completely new apparatus in which the three quantities measured were the e.m.fs. in two equal coils at right angles to one another set at 45° on either side of the geographical direction, and, as a third observation, the phase angle between these e.m.fs., *i.e.*, the sides OA, OB and the angle AOB in fig. 2. The set also incorporated several refinements not found in the earlier ones. Since the completion of this set in June last, it has been found to give a very great improvement in results, and all further deductions have been based on these three observations.

It is now possible to discuss the evidence in favour of plane polarisation. In the first place, while from the purely mathematical standpoint we have only three equations for four unknowns, there is the further physical criterion that any solution must be real and physically possible, and probably possess reasonable continuity in time. In the present case this requires that the ratio of the actual intensity of the down-coming wave to its theoretical value calculated on the assumptions that the layer is a perfect reflector and the intervening medium a perfect dielectric, must always be less than unity. (Incidentally this ratio is in future referred to as the "equivalent reflection coefficient" of the layer.)

Any assumption which consistently gives results satisfying this condition is at any rate worthy of further investigation. There is, however, more direct evidence. Examination of the observations taken in a considerable number of sunrise and sunset runs shows that the phase angle between the e.m.fs. in the two coils as a general rule changes very much more slowly than their intensities. In one particular case, in the course of 6 minutes the e.m.fs. in the two coils fell to half their initial values while the phase angle between them remained constant at 35°. Examination of the resulting vector diagram shows that if in this case the polarisation had been circular it would have been necessary for the two vectors H_1 and H_1^1 to go through in this period two distinct and elaborate intensity cycles depending not only on themselves but also on their momentary phase and intensity ratios to H_0, actually reversing their direction of rotation relative to H_0 during the period.

On the other hand, the result can at once be interpreted numerically on the assumption of plane polarisation, by a change of 50° in the angle of rotation of this plane, without even a change in the value of the reflection coefficient. Of these two alternatives the second is so very much simpler that there seems every justification for adopting it. The particular case quoted was rather a marked example; but the comparatively slow change of phase is a definite characteristic of these runs; in fact many cases have occurred in which the whole sunrise cycle has been completed with the e.m.fs. in the two coils in phase throughout, and some results are definitely insoluble unless the vectors H_1 and H_1^1 are in phase or nearly so. Hence in the remainder of this paper plane polarisation will be assumed, but it should be noted that there is no intention of extending it by analogy over any range of wave lengths not covered by these actual experiments.

From the diagram in fig. 2 the quantities, η, H^1, and the phase angle between H_0 and H^1 can easily be calculated.

Description of the Apparatus.

Briefly the apparatus consists of two coils about 5 feet square mounted at right angles to one another about a common vertical axis, the whole being free to rotate. Each coil is wound with 54 turns of wire and is separately tuned and its circuit completed through one of the fixed coils of a radio-goniometer. This is merely two small coaxial coils mounted at right angles to one another in the exact position of zero mutual inductance, and another smaller coil, also coaxial with them, but free to rotate about its axis and fitted with a scale and pointer. The whole of the goniometer is heavily screened in a sheet-iron box. The search coil is connected to a third tuned circuit, also screened, the condenser terminals of which are joined to the grid and filament of the first valve of an amplifier. The amplifier is of the resistance-capacity type, which, though of slightly lower efficiency than some other forms, will work over a considerable wave-length range without re-adjustment. The output from the last valve is connected to a crystal in series with a small string galvanometer fitted with a microscope and eye-piece scale, on which the actual deflections caused by the incoming Morse signals can be read directly. For calibration the input terminals of the amplifier are connected to a source of known and adjustable e.m.f. of the same frequency as the incoming signal. From this reading and the known circuit constants of the set, the incoming e.m.f. can be calculated. For determination of the phase angle use is made of the fact that, if the two e.m.fs. in the main coils are in phase, the currents induced in them will give rise to a linear resultant in the goniometer; so that, by rotating the search coil to a position at right angles to this, a clear sharp zero will be obtained on the string galvanometer. If, however, they are not in phase, their currents will give rise to an elliptical field in the goniometer, and no absolute minimum will be obtainable on the search coil. But by slightly de-tuning one of the coils, so as to alter the phase relation between the e.m.f. and the current in it, the sharp minimum can again be obtained; and the reading of the amount of de-tuning required (which is done by means of a calibrated condenser) together with the knowledge of the circuit constants, gives the value of the phase angle between the two e.m.fs. This has been found to work very satisfactorily except in the two extreme cases when the original phase difference is nearly 90° (which hardly ever happens), or when one of the e.m.fs. is so much smaller than the other that the exact position of de-tuning balance is uncertain. This has occurred occasionally, but in this case the abnormal polarisation is so great that fairly accurate results can be obtained even from rough observations.

Experimental Results.

After some consideration and several experiments it was finally decided to confine the observations for the time being to a single transmitting station.

The reason for this was that it soon became clear that there were very distinct day-to-day variations; so that to begin by observations on several stations would merely have increased the number of independent variables without adding any generalisation of any value.

Moreover when abnormal polarisation is present, unambiguous values of the quantities measured cannot be obtained from a single observation. For the effect of a 180° change of phase between H_0 and H^1 with unaltered polarisation is identical with that of a 180° change of rotation of plane of polarisation with unaltered phase, although the intermediate stages are altogether different. Hence it is strictly necessary to take a series of observations beginning or ending with a time at which the state of polarisation is roughly known. Now all experience shows that the rotation is zero during the day (in summer at any rate), and so for definite results a complete sunset or sunrise cycle is required. It is, however, possible to dispense with this occasionally after the general characteristics of a particular transmission have become known by experience. Thus in general it is impossible to take simultaneous comparative observations on more than one station without duplication of apparatus.

The transmission first adopted was that from St. Assise (near Paris) on 14,350 metres. This has the advantage that it is at a distance from Slough which exhibits the phenomena under consideration in a very marked way. Also its transmission is very good and steady, and it is working most hours of the day and night. The author is under a particular debt of gratitude to the Compagnie Radio-France, who have kindly supplied him with a great deal of information without which he would have had great difficulty in making the necessary calculations.

The first observations were made during a period covering the solar eclipse of June 29, 1927, as it was thought that useful information might be obtained from the combination of these and the eclipse observations. As an example of the results obtained, fig. 3 gives in graphical form the actual intensities measured on the two coils during sunset on the day preceding the eclipse, from which it will be seen that the variations are considerable and reasonably continuous; in fact it seems quite certain that there are no large fluctuations with periods measurable in seconds. The string in the galvanometer indicating the Morse signals comes to a clear and definite position for each dot and dash

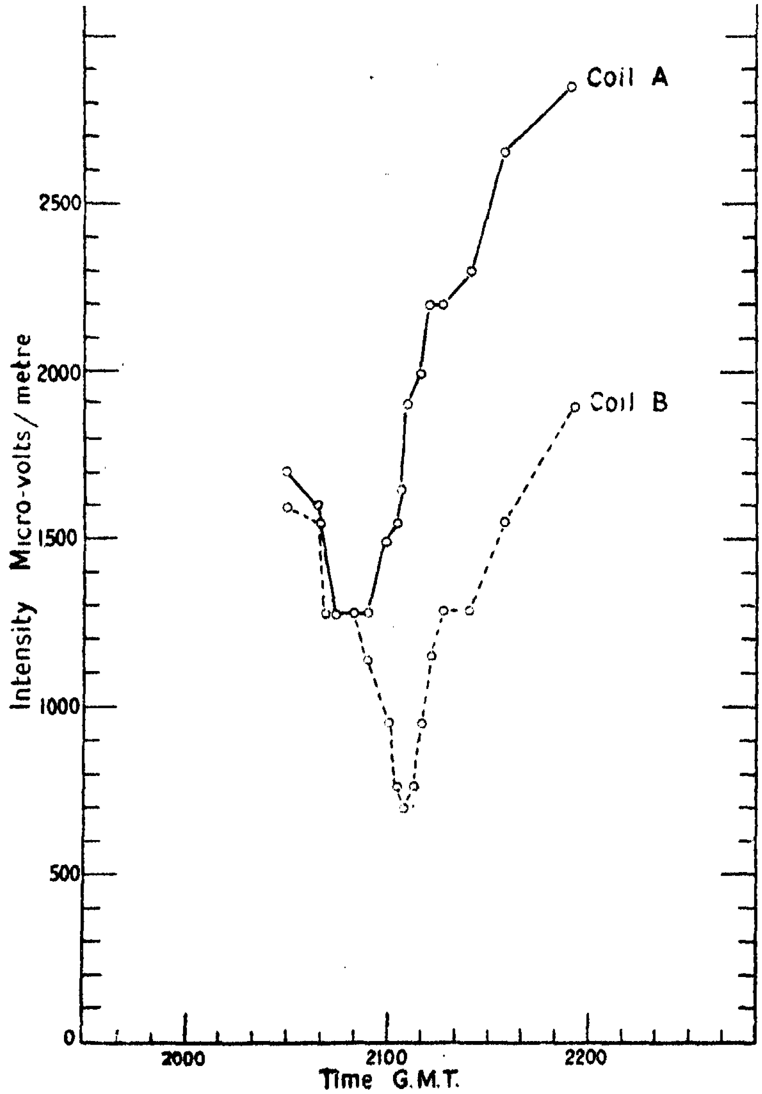


Fig. 3.—Curves showing the Actual Readings obtained on the Measuring Set, from which the Results in fig. 4 are calculated.

(except in the case of high-speed automatic sending), and it is this position which slowly fluctuates as shown in the curves.

The analysis of these curves is given in fig. 4, and many sunrise and sunset curves have been obtained since that date, of which fig. 5 is a recent specimen.

Individually they show certain variations, for a full examination of which it will be necessary to accumulate more data covering all seasons of the year. Such observations, however, have to be started at least one hour before sunrise, and sometimes continued till after sunset. Owing to the use of a null method for the phase measurement it cannot be done automatically, and the attendance of an experimenter is required continuously; with the result that physical

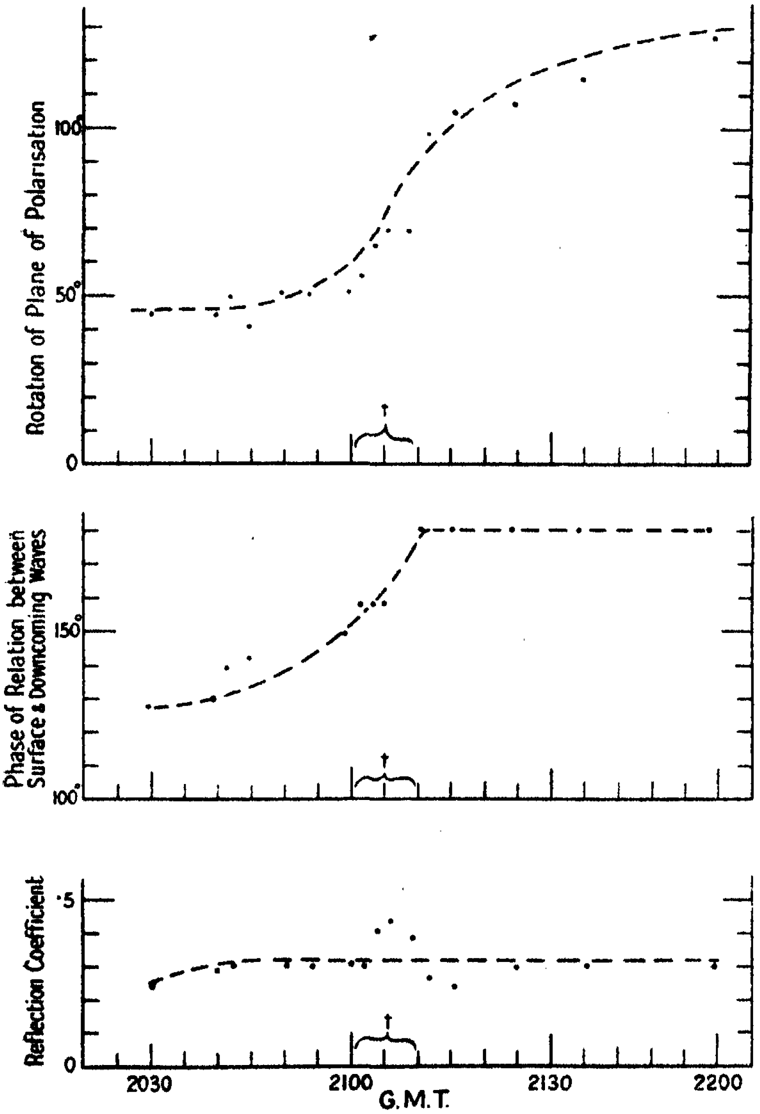


FIG. 4.—Analysis of Sunset Run on St. Assise, taken at Slough, June 28, 1927.

limitations lead to a comparatively slow accumulation of results. But all those so far obtained show a certain number of distinct and well-marked similarities, and it is only with the discussion of these that the present paper is concerned.

General Deductions.

The three outstanding physical quantities arising out of the results can be grouped as follows :—

(*a*) Reflection coefficient.
(*b*) Rotation of plane of polarisation.
(*c*) Phase relation between surface and down-coming waves.

(*a*) *Reflection Coefficient.*—The reflection coefficient has been defined as the ratio of the actual down-coming intensity to the value it would have if reflection at the layer were perfect and the air a perfect dielectric. It therefore includes absorption losses in the intervening medium as well as actual reflection losses. One of the most striking of its features is its great constancy in spite of the large observed variations from which it is calculated. For example, the two following observations taken entirely at random from a single run both gave the same value (0·5) for the reflection coefficient :—

Intensity in coil A.	Intensity in coil B.	Phase angle. °
800	1100	+26
3300	2200	−17

Similar cases are of continual occurrence, and this in itself is taken as a further, though indirect, justification of the assumption of plane polarisation. The most usual night value is from 0·3 to 0·4 with occasional ranges from 0·2 to 0·6. The day value (which is somewhat difficult to determine) is usually 0·1 to 0·15, though there are signs that it is occasionally zero. No attempts have yet been made to correlate it with other phenomena, and at present its value seems to be entirely a matter of chance.

Another feature is that the rise from the daytime value at sunset and the corresponding fall at sunrise take place about the times of sunset and sunrise on the earth's surface or even before and after them respectively. The exact times of these changes and the daytime value itself are rather difficult to determine accurately, as, when the reflection coefficient and abnormal polarisation are both small, high accuracy is required to obtain soluble equations. The fact, however, remains definite ; and the full implication of it will be discussed later.

(*b*) *Rotation of the Plane of Polarisation.*—All the curves, specimens of which are given in fig. 5, show distinctly that the change in the rotation of the plane of polarisation is in two parts, each fairly abrupt; and that the first of these parts is over in from 10 to 12 minutes. The exact time of its commencement can often be determined more accurately at sunrise than at sunset, since the reflection coefficient is usually higher and the effect on the intensities more

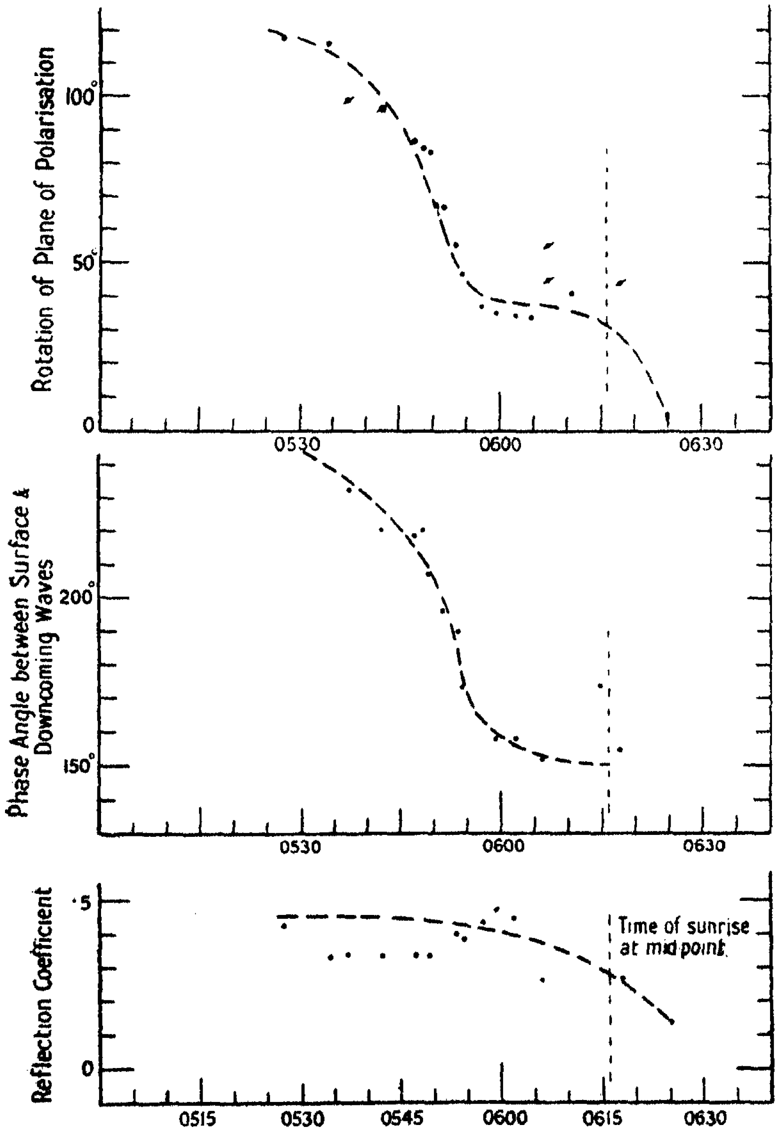


FIG. 5.—Analysis of Sunset Run on St. Assise taken at Slough, October 25, 1927.

marked. Under favourable conditions the commencement of the rapid change (see fig. 7) can be determined to half a minute, and so sunset runs are now seldom taken. From this curve it will be seen that while there is a gradual drop for 15 minutes or so preceding the critical time (0547), at this instant there is an unmistakable change in rate of decrease in intensity. Observations were actually taken at 1 minute intervals throughout the run; but prior to 0547 they are not all shown on the diagram as they lay uniformly on the curve drawn. As will be noticed this change is shown much more prominently in fig. 7 than in fig. 5. The reason for this is that in fig. 5 the angle of rotation is derived from its tangent, and when the former is near 90° its tangent changes much more rapidly than the angle itself. (In all these diagrams except fig. 7 the actual curves shown are only intended as rough indications of the forms obtained.)

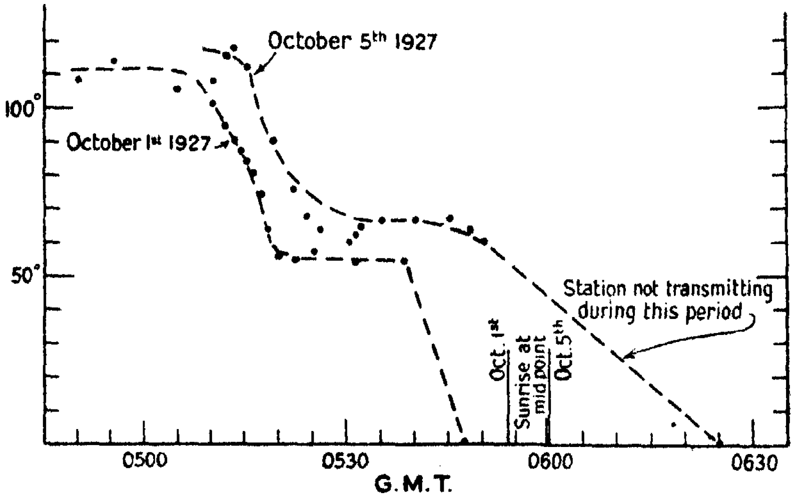


FIG. 6.—*Typical Polarisation Curves for Sunset Period.*

In connection with this time of commencement there are some interesting deductions. At the time of writing (October) the effect is occurring about 45 minutes before sunrise at the point midway between transmitter and receiver. On examining previous records for June, the time interval is found to be 55 to 60 minutes; the lower accuracy being merely due to the fact that at that time its importance was not fully appreciated and so its time of incidence was not so accurately noted. By calculation from the actual readings it is found that owing to the change in the sun's declination between these periods, they all correspond to a position of the sun 7° below the horizon at the mid-point referred to above. We can thus assume that the ionisation of the upper layer by sunlight starts when the sun is 7° below the horizon at the point on the earth's surface vertically below. Now this means that the sun is just rising (neglecting refraction) at a point in this vertical line 47 km. above the earth's surface.

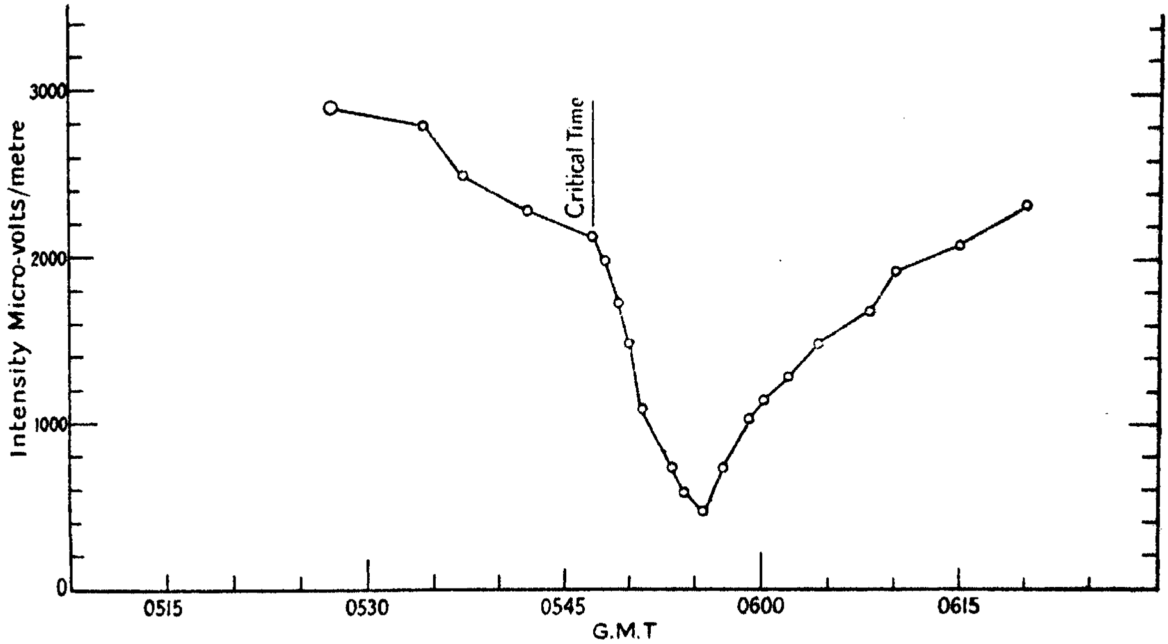


Fig. 7.—Graph of Actual Readings taken on October 25, 1927, showing critical period at 0547 G.M.T.

This fixes a lower limit for the height of the layer. The actual height must be greater for the following reason. Such a ray, being tangential to the earth's surface, would pass twice through the thickness of the earth's atmosphere, and so would have its ionising power completely destroyed. It is not until it has risen sufficiently above the lower atmosphere to retain this power that it can begin to have any effect on the layer. So little is known about the absorbing power of the upper atmosphere that this cannot be allowed for; and it seems more reasonable to argue that, as other purely radio experiments give the height of the layer as 75–85 km. the absorbing power of the atmosphere for ultra-violet radiations persists up to a height of 30 to 40 km. above the earth's surface, after which it ceases fairly abruptly. Further experience has suggested that this absorption of the sun's ionising radiation by a medium through which it has passed may be an important factor in several radio phenomena.

During the first change the rotation usually drops to about 30 per cent. of its night value and remains constant for a time until the second abrupt change occurs and it again drops, this time to zero (summer period). The exact time of this second change is more difficult to determine for the reasons given above; but it appears to be about sunrise on the earth's surface at the mid-point. Now this time, the sun's rays being then tangential to the layer, is, of course, a transition period between the time when the layer is illuminated from below and that (normal daytime) when it is illuminated from above. It is conceivable that these two methods of illumination cause different ionic gradients, and so account for the change; but of course this has, at present at any rate, no proof.

The actual amount of rotation occurring varies daily and a discussion of it is reserved for another paper.

Observations continued throughout the night have shown that after the rapid sunset changes there may be slight further changes for an hour or so, after which a steady state is reached which persists all night. Actually in the middle of the night the intensities seem to be almost as steady as in the day, but of course the values are usually very different. (This applies primarily to the summer months, extended data for the winter period not having been obtained at present.)

It should be noted that there is always a considerable difference in time between the commencement of the abnormal polarisation and of the change of reflection coefficient referred to in the preceding section. While the former appears abruptly on an average about 45 minutes after sunset and before sunrise, the latter is less definite, but is certainly much nearer to the actual times of sunrise and sunset on the earth's surface at the mid-point. Now at these latter times the layer itself is in full sunlight, and has been or will remain so for at least another half-hour, so that there is a distinct suggestion that an appreciable amount of the energy loss of the reflected wave occurs in the medium between the ground and the active layer rather than in the layer itself. This of course is no new idea, but rather a confirmation of an old one by a new method. If this is the case the expression "reflection coefficient" will have a very limited and special meaning. For this reason it was on its introduction specially referred to as the "equivalent reflection coefficient," and defined on p. 451 in such a way as to include this, for it will in this case be not only a function of the layer constants but will also contain a distance attenuation factor as well.

(c) *Phase Relation between Surface and Down-coming Waves.*—One of the most unexpected results of these analyses was the small change in the phase relation between these two waves during the sunset and sunrise periods due to change of equivalent path length. Previously the "V" curve obtained for the intensity during the sunset cycle had been explained as entirely due to this, and a consequent change of layer height of 10 to 15 km. inferred therefrom.

It now appears, however, that this effect was largely caused by abnormal polarisation, and that the actual change of equivalent path length is very much smaller, corresponding only to a change of 1·5 to 2·0 km. at the utmost in the most marked cases. Frequently throughout the whole of a sunrise cycle no visible phase change occurs, and in one particular case recently the e.m.fs. in the two coils remained in phase for nearly 24 hours, though during that period each of them went through cycles of intensity of the order of 5 to 1 in magnitude and also changed their relative values by considerable amounts. It has always been felt that the idea of equivalent height had but little physical reality, and in view of these results one is inclined to argue that no real mechanical rising of the layer occurs, but merely a change in its electrical and physical state, and that the phase change effects are secondary effects due to the alteration of the value for the reduced path of the ray in the layer as its electrical constants change. Examination of the results so far obtained does not seem to show any connection between this phase change and the other factors in the transmission, but for this purpose a larger number of observations will probably be required than are at present available.

Before finally summing up it may be well to give a short discussion of probable errors and their effect on the results.

Quantitative radio work suffers from many limitations. On the purely technical side the instrumental accuracy is inevitably much lower at these high frequencies than it is at those in commercial use, and in addition, when working on ordinary station routines (as is almost essential in long-wave work) neither the time nor the nature of the transmission are under the observer's control. It frequently happens that in the middle of a vital series of observations the station shuts down. No organisation could remedy this, as experiments are so frequently dictated by the needs of the moment. In addition to this, all the calculations involve the use of vector diagrams, and cases are bound to occur in which the triangles involved are of such degenerate form that the accuracy of solution is low.

Tests have shown that under favourable conditions relative measurements can be obtained to an accuracy of 2 per cent., and it is on these that the shape of the vector triangle usually depends. Due to both internal and external causes, absolute accuracy is lower, probably 5 per cent to 10 per cent., but in these calculations the chief effect is to change the reflection coefficient by a similar percentage, and at times to make a somewhat greater change in the angle of rotation. Neither of these changes has any serious effect on the present work.

Summary.

It will be seen that the present paper is rather of an introductory nature to a subject which has as yet received but little systematic study. Routine observations are now proceeding in order to obtain a greater quantity of data, and to study certain special effects. Even from the present results, however, certain general characteristics have shown themselves in an indisputable manner, and it is with the discussion of these that the paper is chiefly concerned. In summary form its main conclusions are given below.

(*a*) During the hours of darkness long wave intensities are moderately steady, but show considerable abnormal polarisation.

(*b*) At a time before sunrise corresponding to the sun's position of 7° below the horizon at a point on the earth's surface mid-way between the transmitter and receiver, an abrupt change lasting about 12 minutes occurs, during which the polarisation becomes much less.

(*c*) A similar effect with increase of abnormal polarisation occurs at sunset, but its period is sometimes rather less definite and slightly longer.

(*d*) At sunrise the value of the reflection coefficient which has remained at about its night value throughout the period (*b*), falls gradually to its day value which in summer is usually about 30 per cent. of its night value (but appears to be zero occasionally), and during this period the remainder of the abnormal polarisation disappears.

(*e*) The evidence points to a much smaller change in equivalent height of the layer than has been assumed up to the present; it also suggests that a considerable proportion of the loss of energy in the indirect ray may occur below the actively refracting part of the layer rather than in it.

Note.—The general features of these deductions are unaffected by the particular form of polarisation assumed, but it must again be emphasised that they only apply to long waves (> 10,000 metres) and to distances of transmission between 400 and 1000 km.

General Observations.

In view of the increasing use and importance of short waves it may be thought that the time is now past for the study of long-wave phenomena. While it must be admitted that no theory or study of propagation is complete which does not deal with short waves, it must be pointed out that the critical study of the subject by other than statistical methods has not yet advanced very far. It seems advisable therefore to study the subject under the conditions in which it is most amenable to quantitative analysis, rather than necessarily under the practical conditions of everyday use; and for this purpose there is no doubt that long waves possess very great advantages, as the instrumental problems which arise are very much less, and experience has shown that the variations which occur are slower and less erratic. Logically therefore it seems that the long wave problems should be studied in detail first, the extension to shorter waves following later.

Below 400 km. distance the conditions of propagation appear to be entirely different, and are being made the subject of a separate investigation. Their salient characteristics appear to be (1) reflection coefficient usually zero, but may be as high as 0·4; (2) whenever a refracted wave is received, by day or night, it is abnormally polarised, and no marked change occurs during the sunrise and sunset periods; (3) the few measurements so far obtained suggest an effective vertical height of penetration of about 150 km.

With regard to Prof. Appleton's recent letter to 'Nature'* as to the existence of another higher layer operative in the middle of the night, no definite evidence has yet been obtained. Although the height of the present layer assumed (80 km.) is too low for direct determination by the aerial-coil method for stations at the distances apart employed in the present paper, this would not be the case for a height of 150 km. Tests of this method have been made during the night; but unfortunately at this time the usual state of polarisation of the signals is a rotation of about 90° in which case the critical term in the equation is nearly zero, and a solution is paractically unobtainable.

On the other hand, if the signals penetrate to such a layer during the night and fall back to the normal one during the day, one would expect enormous variations at the period of change, if not almost a discontinuity; and in none of the observations taken so far has any such effect been noted, although the signal is under absolutely continuous observation during the periods of watch.

At distances above 1000 km. the scientific side of the problem becomes much more complex. Multiple reflection may occur, and signs of it at this distance of 1000 km. have recently been detected and are now being investigated. Abnormal polarisation may also be present; though it cannot in general be detected or measured owing to the absence of an appreciable ground wave and the high angle of incidence of the refracted ray, which make the ordinary polarisation measuring methods unworkable. It seems likely, however, that it should always be considered as a possible factor in any theory of long distance propagation, as, if present, small changes in it may have very large effect on the signal intensity; a wave with its electric vector horizontal having no effect whatever on normal receiving apparatus.

---

* 'Nature,' vol. 120 p. 330 (1927).

This work described was carried out as part of the programme of the Radio Research Board and is published by permission of the Department of Scientific and Industrial Research.

Finally, I must express my sincere gratitude to my colleage Mr. Naismith, both for his suggestion and carrying out of the original experiment which led to the work being undertaken; and also for his share in the construction of the apparatus and in the taking of the observations; involving as this does protracted and delicate observations over periods of several hours at all times of day and night.

Gaseous Combustion at High Pressures.—Part X. The Co-volume Corrections, Maximum Temperatures and Dissociation of Steam and Carbon Dioxide in Explosions.

By Dudley M. Newitt, Ph.D., D.I.C.

(Communicated by Prof. W. A. Bone, F.R.S.—Received April 4, 1928.)

Introduction.

During the researches upon high-pressure explosions of carbonic oxide-air, hydrogen-air, etc., mixtures, which have been described in the previous papers of this series, a mass of data has been accumulated relating to the influence of density and temperature upon the internal energy of gases and the dissociation of steam and carbon dioxide. Some time ago, at Prof. Bone's request, the author undertook a systematic survey of the data in question, and the present paper summarises some of the principal results thereof, which it is hoped will throw light upon problems interesting alike to chemists, physicists and internal-combustion engineers.

The explosion method affords the only means known at present of determining the internal energies of gases at very high temperatures, and it has been used for this purpose for upwards of 50 years. Although by no means without difficulties, arising from uncertainties of some of the assumptions upon which it is based, yet, for want of a better, its results have been generally accepted as being at least provisionally valuable. Amongst the more recent investigations which have attracted attention in this connection should be mentioned those of Pier, Bjerrum, Siegel and Fenning, all of whom worked at low or medium pressures.

The experiments upon which the present investigation is based have included hydrogen- or carbonic oxide-air, etc., explosions over a much wider range of initial pressures (namely, 3 to 175 atmospheres) than any previously; and new information may be deduced from them as to the effects of density and temperature upon both the internal energies of gaseous media and the dissociation of steam and carbon dioxide.

Whilst the experiments in question were in progress it was possible, on certain provisional assumptions which seemed reasonable, to estimate approximately the maximum temperatures actually attained in the various explosions as well as the dissociation of steam or carbon dioxide thereat; but not until the accumulated data could be reviewed as a whole, and some additional experiments also made to elucidate dissociation factors, could any final figures be deduced. For the purpose of this paper, such a revision of the previous estimates of maximum temperatures and dissociation has been carried out; and the figures now given may be regarded as replacing the previous ones, which will be indicated in the text. From these it will be seen that, speaking generally, whilst the revised figures for the maximum temperatures attained in theoretical hydrogen-air and carbonic oxide-air explosions do not differ materially from those previously estimated in cases where the initial pressure did not exceed 25 atmospheres, at all higher pressures the newer estimates are lower than the old, and the more so as the initial pressure increases. This circumstance is due to the fact that a co-volume correction is now made in respect of the maximum pressure obtained in each explosion.

In the previous papers* the opinion had been expressed that, in the explosions referred to, the dissociation of steam at maximum pressure had been "practically negligible," and that that of carbon dioxide had been "fairly constant at about 5 per cent. throughout the whole series of our CO-air explosions." The revised figures now submitted generally confirm these estimates; those for steam show a 2·2 per cent. dissociation at an initial pressure of 3 atmospheres, diminishing regularly to 1·0 per cent. only as the initial pressure was raised to 175 atmospheres. The new CO_2-dissociation figures (which vary only between 4·7 and 5·6 per cent.) all closely agree with the previous estimates.

---

* *E.g.*, Part V, 'Roy. Soc. Proc.,' A, vol. 108, p. 393 (1925).

A remarkable feature of the observed explosion data, to which attention was specially directed at the time, was the fact that the "corrected" P_m/P_i ratios always increased with the initial pressure in an unexpectedly large degree, as the following figures will sufficiently indicate:—

Mixture.	P_m/P_i (corrected).		
	$P_i = 3$.	$P_i = 75$.	$P_i = 150$.
$2H_2 + O_2 + 4N_2$	7·70	8·43	8·82
$2CO + O_2 + 4CO$	8·00	9·48	10·16
$2CO + O_2 + 4N_2$	7·80	9·14	9·76

It was suggested that possibly the explanation might lie "in the increasing opacity of the gaseous medium with pressure to the radiation emitted during the explosion," and undoubtedly this accounts for a substantial part of the effect. It is now shown that about 50 per cent. of it in the case of the CO explosions, and about 75 per cent. of it in the hydrogen explosions, may be ascribed to the co-volumes of the components of the gaseous media in question, and that the correction on this account is not only a function of the initial pressure but also of the maximum temperature attained in the explosion.

Consideration of the Various "Corrections" to be Applied to the Observed Explosion Data.

In carrying out such explosions, the principal observations made are (1) the initial firing pressure (P_i), as indicated by a standard Bourdon gauge; (2) the time taken (t_m) for the attainment of maximum pressure from the commencement of the pressure rise; (3) the maximum pressure (P_m) attained; and (4) the rate of cooling after maximum pressure—the last three being read off from the actual pressure-time record obtained by means of the Petavel manometer and its optically recording adjuncts. For purposes of calculation and comparison of results amongst themselves, certain "corrections" should be applied to the so-observed values of P_i and P_m respectively; these "corrections" are briefly as follows:—

Initial Pressure (P_i).—This should be corrected for deviations from Boyle's Law, because neither term in the PV product is independent of pressure. In cases where P_i does not exceed 10 atmospheres, the correction is so small as to be practically negligible, but at all higher pressures it is not so. Moreover, although the compressibilities of all the individual components of an explosive mixture may be accurately known over a wide pressure-range at room temperature, the failure at high pressure of the Law of Partial Pressures precludes any calculation therefrom of the compressibility of the mixture, the isotherm of which must always be determined by direct measurements at the room temperature. Such direct measurements have always been made in conjunction with our explosion experiments by means of an apparatus specially designed for the purpose; and from such measurements the observed initial pressures have been corrected. Such corrections be it noted include changes in both the pressure and volume terms in the equation of state, and no other "correction" of P_i is necessary.

The actual compressibility factors experimentally determined and used in correcting the observed P_i values, in our experiments, are as follows :—

Mixture.	Observed P_i in atmospheres.					
	10.	25.	50.	90.	120.	150.
$2H_2 + O_2 + 4N_2$	0·9954	0·9938	0·9922	0·9858	0·9836	0·9815
$2CO + O_2 + 4CO$	1·001	1·008	1·020	1·023	1·015	1·006
$2CO + O_2 + 4N_2$	1·002	1·007	1·017	1·016	1·013	0·9974

Maximum Pressure (P_m).—Two corrections may be considered here, namely :—

(*a*) In cases where the attainment of the maximum pressure in an explosion is practically instantaneous—say, as in our case, within a time not exceeding 0·005 second which is actually the "time period" (θ) of the Petavel gauge—it is assumed that the observed P_m is not appreciably affected by any cooling of the gaseous medium by the walls of the explosion vessel. When, however, t_m exceeds 0·005 second, an appropriate "cooling correction" is applied to P_m in respect of the time period by which t_m exceeds 0·005 second. This is always done in comparing the results of two or more experiments with different explosive mixtures, so as to eliminate the influence of observed variations in t_m. For in such cases it is the "P_m corrected" values which are taken into account. When, however, it is desired to calculate from the observed P_m value the mean maximum temperature (T_m) actually attained in a particular explosion, such "cooling correction" is not applied; in such cases, however, it is necessary to allow for "dissociation," and to apply also a co-volume correction.

(*b*) In so calculating T_m at the observed P_m value in a particular explosion, an allowance for "dissociation" comes in in so far as it affects the observed P_m/P_i ratio in respect of the observed degree of dissociation. Thus, for example, supposing, as is the case of a theoretical CO-air mixture, there be a 5 per cent. CO_2-dissociation at P_m; then, instead of the medium at P_m consisting of $2CO_2 + 4N_2$, it consists of $1 \cdot 9\ CO_2 + 0 \cdot 1\ CO + 0 \cdot 05\ O_2 + 4N_2$, and the observed P_m/P_i ratio should be multiplied by 1·157 instead of by 1·166 to allow for the medium at P_m containing 6·05 instead of the 6·00 molecules in an undissociated medium.

(*c*) As in the case of P_i, the observed P_m value in an explosion should be corrected for deviations of the gaseous medium from Boyle's Law* which, of course, cannot be directly measured. It is possible, however, to deduce from first principles on certain very probable assumptions the magnitude of the "correction" which should be applied in a particular case, provided that the maximum temperature attained in the explosion is known approximately. The procedure adopted will now be considered.

Deduction of the Co-volume Correction at the Maximum Temperature in an Explosion.

In applying any one of the numerous equations of state, *e.g.*, that proposed by A. Wohl† in 1914—who applied it satisfactorily to the isotherms of carbon dioxide and of other gases up to pressures of 1000 atmospheres at ordinary temperatures—

$$P = RT/(v - b) - a/[Tv(v - b)] + c/T^2v^3$$

it has generally been assumed that the *b* term (co-volume) is practically independent of temperature, at any rate up to 300° C. The temperature rise in our explosions is of the order of 2000° to 3000° C. and it is doubtful whether the constancy of *b* can be assumed over such a wide range, especially as the viscosity of a gas is known to increase with temperature between 100 and 1000° C. in a manner consistent with the assumption that its molecular diameter diminishes as the temperature rises. On applying Wohl's equation to our explosion results, on the assumption that the second and third term of the equation become negligible at the maximum temperatures concerned, and that the value of b as determined at ordinary temperatures remained constant, the so "corrected" P_m/P_i ratios were found to decrease considerably as the initial pressure rose beyond 10 atmospheres. Thus, for example, in the case of our theoretical hydrogen-air explosions the so "corrected" P_m/P_i values worked out as follows :—

---

* I have to thank Mr. H. T. Tizard, C.B., F.R.S., for a private communication pointing out the magnitude of the co-volume correction in this connection.

† 'Z. Phys. Chem.,' vol. 87, p. 1 (1914).

P_i	3	10	25	50	75	100	125	150	175 Atmospheres.
P_m/P_i	7·70	7·84	7·73	7·51	7·41	7·26	7·02	6·70	6·47

Such a result being manifestly impossible—for it is unthinkable that the maximum temperature attained in such an explosion would diminish as the density of the medium increases—it follows that the assumption as to the invariability of "b" with temperature must be abandoned.

There are two alternative methods of calculating the change in the value of "b" with temperature based upon the view that the space occupied by a molecule of gas is proportional to the third power of the effective molecular diameter.

According to Sutherland's formula, the molecule diameter (σ) varies with the absolute temperature (T) as follows :—

$$\sigma_T{}^2 = \sigma_x \infty^2 (1 + C/T) \tag{1}$$

where σ is the molecular diameter corresponding to T and the constant C is proportional to the work done in separating two molecules, originally in contact, to an infinite distance apart.

The expression as applied to our present problem leads to the following equation for the variation of the co-volume with temperature :—

$$b_T = b_0 \left\{\frac{1 + C/T}{1 + C/T_0}\right\}^{3/2} \tag{2}$$

It has also been found that in a large number of cases the change of viscosity of a gas with temperature is very nearly proportional to the nth power of the absolute temperature T, where n, on the assumption that the repulsive force between two molecules in collision is equal to $\mu\sigma^{-s}$, is given by

$$n = S + 3/2 (S - 1).$$

On this view, the co-volume at a temperature T° is given by

$$bT = b_0 (T/T_0)^{-3/S-1}. \tag{3}$$

The use of equations (2) and (3) for the purpose of correcting the pressure ratios in our explosions lead to results which differ from one another by between 1 and 2 per cent. at the higher pressures. It has therefore been thought advisable to employ both methods, and to compare the values both of the P_m/P_i ratios and of the maximum temperatures so obtained.

The following values for the constants in equations (2) and (3) have been employed* :—

Gas.	C.	S.
Steam	650	4·6
Carbon dioxide	274	5·2
Carbonic oxide and nitrogen	118	9·3

and the value of b_T for the gaseous products obtained in the explosions at the maximum temperature calculated therefrom.

Values of b_T calculated by means of equations (2) and (3).

Products from explosions of—		Value of b_T at temperatures (abs.).				
		273°.	2500°.	2700°.	2900°.	3100°.
$2H_2 + O_2 + 4N_2$	(2)	0·00160	0·000695	0·000688	0·000681	0·000674
	(3)	0·00160	0·000597	0·000580	0·000560	0·000543
$2CO + O_2 + 4CO$	(2)	0·00170	0·000743	0·000737	0·000733	0·000700
$2CO + O_2 + 4N_2$	(3)	0·00170	0·000623	0·000607	0·000583	0·000567

The corrected values of P_m are obtained by substituting values for b_T in the equation

$$P_m \text{ corrd.} = P_m \text{ obs.} (v - b_T)/v$$

where $v = 1/P_1$.

Employing this equation, the P_m/P_i ratios, corrected in respect of both P_m and P_i for "co-volume," become :—

	P_m/P_i (corrected).								
P_i (atmospheres)	3.	10.	25.	50.	75.	100.	125.	150.	175.
(1) $2H_2+O_2+4N_2$ mixture									
Equation (2)	7·70	7·78	7·91	7·88	8·00	8·04	8·03	7·91	7·90
Equation (3)	7·70	7·85	7·93	7·92	8·05	8·11	8·14	8·05	8·09
(2) $2CO+O_2+4CO$ mixture									
Equation (2)	8·03	8·33	—	8·50	8·87	8·88	8·95	8·95	—
Equation (3)	8·03	8·40	—	8·56	8·99	9·02	9·14	9·19	—
(3) $2CO+O_2+4N_2$ mixture									
Equation (2)	7·13	7·44	7·62	7·74	7·87	8·02	7·96	8·06	8·06
Equation (3)	7·13	7·50	7·54	7·80	7·95	8·13	8·10	8·25	8·26

---

* C. J. Smith, 'Proc. Phys. Soc.,' vol. 34, p. 155 (1922); 'Roy. Soc. Proc.,' A, vol. 105, p. 242 (1924); Jeans 'The Dynamical Theory of Gases,' 4th ed.

In applying the co-volume correction derived from either of the two formulæ in question, preliminary calculations showed that the corrected maximum temperatures in all the explosions at initial pressures between 10 and 175 atmospheres (the only ones that matter) would fall within the limits 2500° and 3100°. Between these two temperature limits the diminution in the co-volume constant b is no more than about 10 per cent.; therefore, in calculating the maximum temperature, first of all an approximate estimate was made of it, using an appropriate co-volume constant between 2700° and 2900°, and such preliminary estimate of the maximum temperature derived always came within 100° of the final value. A final calculation was now made using the co-volume factor corresponding with the so-estimated temperature.

In the cases of mixtures (1) and (2), no "cooling correction" is necessary, because in each case t_m did not exceed 0·005 second; but in the case of the theoretical CO-air mixture (3), where t_m increased from 0·07 second at $P_i = 3$ atmospheres to 0·56 second at $P_i = 175$ atmospheres, a cooling correction must be superposed on the co-volume corrected P_m/P_i values in cases where comparisons with other mixtures are instituted. With such cooling-correction so superposed, the completely corrected P_m/P_i values for the $2CO + O_2 + 4N_2$ mixture become :—

	P_m/P_i corrected for cooling and co-volume.								
P_i (atmospheres)	3.	10.	25.	50.	75.	100.	125.	150.	175.
Equation (2)	7·80	7·94	8·16	8·35	8·67	8·72	8·64	8·75	8·95
Equation (3)	7·80	7·95	8·17	8·38	8·72	8·79	8·73	8·88	9·10

It will thus be seen that, in the case of the theoretical hydrogen-air explosions, from 70 to 80 per cent. of the observed great increase in the P_m/P_i ratios as the initial firing pressure was progressively raised from 3 to 175 atmospheres may be accounted for by the probable increase in the co-volume factor b with temperature and pressure. In the case of both the theoretical CO-air and the $2CO + O_2 + 4CO$ explosions, however, only about half of the observed increase could be so accounted for. The remainder is probably due to an increase in the opacity of the medium with pressure, a supposition for which the research as a whole has afforded considerable support.

The Dissociation of Steam and Carbon Dioxide at the High Temperatures and Pressures attained in the Explosions.

Having corrected the observed P_m/P_i ratio for deviations from Boyle's Law in the way described in the previous paragraphs, it remained to determine the "degree of dissociation" of steam or carbon dioxide at the maximum pressure in each explosion before the maximum explosion temperature could finally be deduced. For, if P_i and P_m are the respective initial and maximum pressures, after correction in each case for deviation from Boyle's Law but not for "cooling," e the ratio of the number of molecules before firing and at P_m, respectively, and V the volume of the explosion vessel, then

$$T_m = T_i \times P_m V \,.\, e/P_i V,$$

when T_m = the actual mean maximum temperature (abs.) attained in the explosion and T_i = the initial firing temperature (abs.). And, since V does not sensibly alter during the explosion, it may be said that

$$T_m = T_i \times P_m e/P_i.$$

Some General Considerations.—Before proceeding, however, to show how the factor e has been determined, it seems desirable to say something more about the physical condition of the gaseous medium during our explosions, all of which (it may be recalled) were carried out in spherical steel bombs with an explosion chamber of 240 c.c. capacity, ignition being effected by the electric glowing of a fine platinum wire at a position near the walls of the chamber opposite to the head of the Petavel recording manometer. It may be here said that experiments have shown that, so far as the observed t_m and P_m values are concerned, it matters not whether ignition of the charge is effected as described or at a position right up against the head of the Petavel gauge. At such high firing pressures as have been used in most of our experiments, "central-ignition" is impracticable for reasons which need not be discussed here; but, had it been practicable, the observed t_m's would have been uniformly about two-thirds those obtained with the firing position actually employed. Obviously the t_m observed in any particular case must depend upon the shape and dimension of the explosion chamber, and to some extent also upon the firing position in it; but, so long as these factors remain the same throughout a series of experiments, the observed t_m's are all comparable among themselves.

In all such explosions, a measurable time is required after ignition for the flame to spread throughout the gaseous medium and for the observed maximum pressure to be attained. Strictly speaking, probably at no time during an explosion is the gaseous medium absolutely homogeneous throughout, especially as to temperature; nor can the "cooling effect" of the vessel walls ever be entirely eliminated. In general, however, the denser the medium exploded, the shorter is t_m, and the less the said "cooling effect." Indeed, the only exception to this rule found so far at initial pressures between 10 and 175 atmospheres is afforded by our CO-air mixtures, where t_m rapidly increases with the initial pressure, owing to "nitrogen-activation" coming in. And it is the speeding up of t_m and the minimising of "cooling effects," as well as the high medium density itself which make high-pressure explosions so valuable a means of exploring gaseous interactions at high temperatures and densities.

It should always be borne in mind that, in speaking of the "maximum" temperature or pressure attained in any explosion in a closed vessel, the *mean* condition of the gaseous medium as a whole at the moment when the manometer records the maximum effect is meant; hence, the nearer the medium to absolute homogeneity at such moment, the nearer will the observed mean maximum temperature or pressure approximate to the true maximum explosion condition. In a bomb of infinite size no such approximation to homogeneity could ever be postulated; and the employment of explosion chambers of dimensions greater than are necessary to minimise the "cooling effects" of the walls would seem to be disadvantageous. In the case of our bombs, however, the moderate size of the explosion chamber (240 c.c.), whilst sufficient (at all initial pressures above about 10 atmospheres) satisfactorily to minimise "cooling effects," ensured a high degree of homogeneity of the medium at the moment of maximum pressure.

Attention has already been directed to the absence of any sign of what has been termed "after-burning,"*—as distinct from "dissociation" effects at maximum pressure—in any of our explosions where the initial pressure exceeded 10 atmospheres, although in explosions at lower initial pressures such signs were distinctly present. Indeed, speaking generally, the characters of the "differential" cooling curves after the attainment of maximum pressure in explosions where $P_i = 10$ atmospheres seem decisive against the supposition of "after burning," as distinct from "dissociation." In this connection may be recalled the highly significant circumstance that in all the $2CO + O_2 + 4CO$ explosions at initial pressures of 50 atmospheres and upwards, the mean rate of cooling of the medium during the 0·5 second immediately after P_m was always proportional to something between the third and fourth power (increasing with the density of the medium) of the absolute temperature, a result hardly compatible with any "after burning" in such cases.

---

* 'Roy. Soc. Proc.,' A, vol. 105, p. 417 (1924).

Experimental Determination of Dissociation of Steam and Carbon Dioxide in Theoretical H_2*-air and* CO*-air Explosions.*

Dissociation of Carbon Dioxide.—A large number of experiments have been made with the object of determining the degrees of dissociation of carbon dioxide at the maximum temperature and pressures attained in our various theoretical carbonic oxide-air explosions, which covered a very wide range of initial pressures. Two methods were available, namely, those used by N. Bjerrum and by Tizard and Fenning, respectively, and both have been employed.

In Bjerrum's method, two mixtures are found at each required temperature, the one containing a large excess of the combustible gas and the other a corresponding excess of the nitrogen, such that on explosion the same maximum pressures are attained. Assuming that sufficient excess of combustible gas is present in the first mixture to suppress dissociation completely, the composition of the products of combustion of the second mixture may be calculated, and hence the percentage dissociation of the steam or carbon dioxide. Two further assumptions are made, namely (1) that the heat losses during explosion are the same for the two mixtures, and (2) that the specific heats of the two diatomic gases are identical at the high temperatures attained. The assumption as to equality of heat loss is only true when the explosion times of the two mixtures happen to be the same; consequently in finding corresponding mixtures the effect of differences in heat loss must be taken into account. The following example of the method, as applied to CO-air explosions, will perhaps suffice.

It had been previously found that the observed t_m and P_m/P_i ratio for a $2CO + O_2 + 4N_2$ explosion at an initial pressure of 10 atmospheres were 0·10 second and 7·5 respectively.* And, by interpolation from the results of the following further experiments upon $2CO + O_2 + x\,CO$ mixtures at the same initial pressure (see Table I below), the "corresponding" carbonic oxide-oxygen mixture was found to contain 13·3 per cent. of oxygen and to have a $t_m = 0{\cdot}05$ second.

The difference between the "cooling losses" up to P_m in the two cases was estimated from the respective cooling curves as being 2 per cent.; therefore, the corresponding CO-O_2 mixture having the same cooling loss as the theoretical CO-air mixture at the said initial pressure is one containing 13·5 per cent. of oxygen, and giving on explosion a P_m/P_i value = 7·65. The carbon-dioxide content of the exploded products of the two mixtures at the maximum pressure will be identical, so that at P_m in the CO-air explosion :—

---

* *Vide* Part V (*loc. cit.*), p. 406.

$$0{\cdot}286\ CO + 0{\cdot}143\ O_2 + 0{\cdot}571\ N_2 = 0{\cdot}270\ CO_2 + 0{\cdot}016\ CO + 0{\cdot}008\ O_2 + 0{\cdot}571\ N_2$$

from which it follows that, the " dissociation constant "

$$K_p = \left(\frac{0{\cdot}016}{0{\cdot}270}\right)^2 \times \frac{0{\cdot}008}{0{\cdot}865} \times 75 = 0{\cdot}0024.$$

Accordingly, the percentage CO_2-dissociation at the observed maximum pressure of 75 atmospheres ($T_m = 2505°$ abs.) in the said theoretical CO-air explosion was 5·6 per cent.

Table I.—Showing Results of $2CO + O_2 + x\,CO$ Explosions at $P_i = 10$ Atmospheres.

Mixture.	O_2-content.	t_m.	P_m/P_i obs.
	per cent.	second.	
$2CO + O_2 + 4CO$	14·3	0·035	8·40
$2CO + O_2 + 5CO$	12·8	0·060	7·18
$2CO + O_2 + 6CO$	11·3	0·110	6·42

In a similar manner, the dissociation-constants of carbon dioxide at the observed maximum pressures and temperatures in our theoretical carbonic-oxide-air explosion at initial pressures of 50, 75, 100 and 125 atmospheres have been ascertained, with results as shown in the following table :—

Table II.

P_i atmospheres.	Percentage O_2-content of corresponding CO-O_2 mixtures.	t_m	T_m	K_p
		seconds.	° abs.	
10	13·50	0·05	2505	0·0024
50	13·61	0·05	2610	0·0083
75	13·62	0·06	2660	0·0120
100	13·53	0·05	2715	0·0230
125	13·57	0·05	2710	0·0250

The results so obtained by Bjerrum's method have been compared with those experimentally obtained by the method employed by Tizard and Fenning, which consists in progressively increasing the amount of combustible gas in a $2(1+a)CO + O_2 + bN_2$ mixture (the ratio N_2/O_2 being kept constant) until a value for a is found for which the maximum pressure attained on explosion is greatest. At low initial pressures, however, the agreement between the results of the two methods was not good, a circumstance probably due to the fact that, owing to the rather large surface volume ratio of our bomb, it is only at higher initial pressures that the two methods would be really comparable in our case. Also, the undoubted occurrence of a little "after burning" in our theoretical CO-air explosions at initial pressure up to (but not exceeding) 10 atmospheres would render the results of the two methods incomparable at such low pressures. The agreement between the results found by the two methods was, however, fair at an initial pressure of 50, and close at 125 atmospheres; but for the purpose of the calculations included in this paper, those obtained by the first (Bjerrum's) method have been adopted as the more reliable for our conditions.

The value of K_p so found for CO_2-dissociation over the temperature range 2400° to 2800° *abs.* may be represented by the empirical equation :—

$$\log K_p = 7\cdot500 - 25{,}000/T.$$

Such K_p values, it may be noted, fall below those obtained by Bjerrum, but above those of Tizard and Fenning. Thus at 2500° abs., whereas my values for $K_p = 0\cdot0024$, according to Bjerrum it is $0\cdot0040$, and to Tizard and Fenning $= 0\cdot0011$; again, whereas at a temperature of 3200° abs. Tizard and Fenning obtained $K_p = 0\cdot35$, my value is $0\cdot49$, that of Bjerrum at 3116° being $0\cdot70$. Therefore, whilst agreeing with Tizard and Fenning's conclusion that Bjerrum's own estimations of CO_2-dissociation in CO-air explosions are probably too high, my experiments would show theirs to be rather too low.

From my so-experimentally determined values for CO_2-dissociation, the following contraction factors e, and maximum explosion temperatures (T_m abs.) have been deduced from the theoretical $2CO + O_2 + 4N_2$ mixture when it was exploded in our experiments at various initial pressures between 3 and 175 atmospheres, the proper allowance having been made in each case in calculating T_m for all deviations from Boyle's Law both at P_i and P_m, respectively.

Table III.—Explosions of $2CO + O_2 + 4N_2$ Mixture.

P_i (atmospheres)	3.	10.	25.	50.	75.	100.	125.	150.	175.
P_m/P_i corrected—									
Equation (2)	7·13	7·44	7·52	7·74	7·87	8·02	7·96	8·06	8·06
Equation (3)	7·13	7·50	7·54	7·80	7·95	8·13	8·10	8·25	8·26
T_m absolute—									
Equation (2)	2385°	2485°	2505°	2590°	2630°	2680°	2665°	2695°	2700°
Equation (3)	2385°	2505°	2525°	2610°	2660°	2715°	2710°	2755°	2760°
Percentage CO_2-dissociation at T_m	5·1	5·5	4·7	4·9	4·9	5·4	5·0	5·6	5·3

Contraction factor *e* 1·155 to 1·158.

Both P_i and P_m are corrected for deviation from Boyle's Law; but not P_m for "cooling," because it is the mean T_m actually attained in the explosion which is required.

It will be seen how very well the values for CO_2-dissociation so found agree with the estimate of a CO_2-dissociation "*of the order of* 5 *per cent.*" in the said explosion previously given in Part V hereof (*loc. cit.*), pp. 410 and 417); the maximum temperatures actually attained in the various explosions as *then* and *now* compare as follows :—

Date.	P_i (atmospheres).								
	3.	10.	25.	50.	75.	100.	125.	150.	175.
	°	°	°	°	°	°	°	°	°
Calculated T_m in 1925	2393	2523	2573	2693	2793	2903	2943	3093	3143
Calculated T_m in 1928	2385	2505	2525	2610	2660	2715	2710	2755	2760

Up to $P_i = 25$ atmospheres the difference between the two sets of figures is not great; it is only at higher initial pressures, where the new "co-volume" correction at P_m comes in with increasing force as the density of the medium increases, that the differences become so.

In our $2CO + O_2 + 4CO$ explosions, the combined effects of excess of combustible gas and high pressure are sufficient to suppress dissociation altogether; also the comparatively short t_m's involve very small cooling corrections. The "corrected" P_m/P_i ratios for deviation from Boyle's Law both at P_i and P_m, and the T_m values calculated therefrom, are shown below :—

Table IV.—Explosions of $2CO + O_2 + 4CO$ Mixtures.

		P_i (atmospheres).						
		3.	10.	50.	75.	100.	125.	150.
P_m/P_i corrected	Equation (2)	8·03	8·33	8·50	8·87	8·88	8·95	8·95
	Equation (3)	8·03	8·40	8·56	8·99	9·02	9·14	9·19
T_m absolute	Equation (2)	2710°	2810°	2865°	2990°	2995°	3020°	3020°
	Equation (3)	2710°	2830°	2880°	3030°	3041°	3080°	3100°

These T_m values are all between 270° and 370° higher than in the corresponding $2CO + O_2 + 4N_2$ explosions (*q.v.*) chiefly on account of the complete suppression of dissociation by the excess of combustible gas.

Dissociation of Steam.—Tizard and Fenning's method has been found quite reliable for deducing experimentally the K_p values for H_2O-dissociation in explosions of the theoretical $2H_2 + O_2 + 4N_2$ mixture at initial pressures of 50 and 125 atmospheres, respectively; but it is not so reliable at 10 atmospheres, for reasons already indicated. The experimental results (with $P_i = 50$ and 125 atmospheres) which need hardly be detailed, are summarised in the curves (fig. 1) showing in each case the relations between P_m/P_i (ordinates) and percentage of hydrogen in the mixtures (abscissæ). Both curves show well

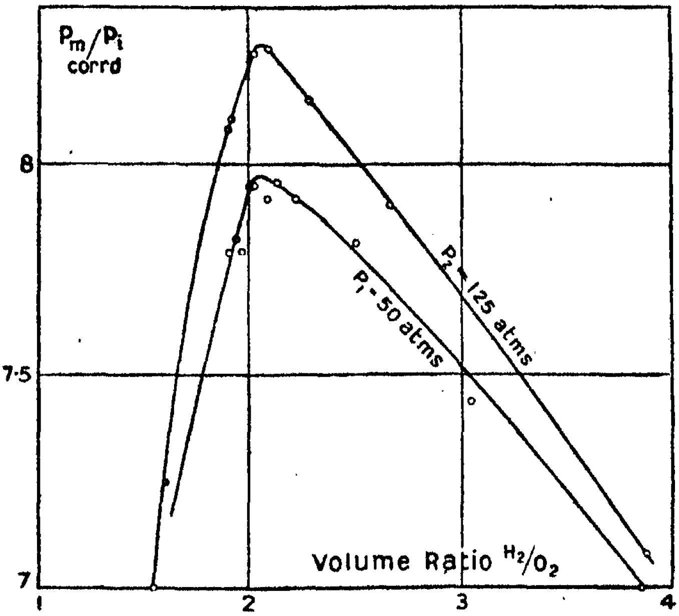


Fig. 1.

defined "maxima," with $t_m = 0 \cdot 005$ second in each case. The experimental details for the mixture giving the maximum P_m/P_i ratio at each initial pressure, are as follows :—

P_i atmospheres.	H_2/O_2 ratio in mixture.	a.	P_m/P_i (corrd.).	T_m abs.	$x = a/b$.	K_p.
50	2·03	0·02	7·95	2682°	0·0053	0·00022
125	2·08	0·02	8·28	2800°	0·0053	0·00060

It may be remarked that the K_p values so ascertained agree very closely with those recently determined by W. Siegel,* who found 0·00022 at 2602° and 0·00053 at 2800° abs.

On the foregoing basis, and correcting both the observed P_i and the P_m values for deviation from Boyle's Law (as already indicated) the following revised data for our theoretical hydrogen-air explosions at the various initial pressures have been deduced :—

Table V.—Explosions of the $2H_2 + O_2 + 4N_2$ Mixtures.

P_i (atmospheres)	3.	10.	25.	50.	75.	100.	125.	150.	175.
P_m/P_i corrected—									
Equation (2)	7·70	7·78	7·91	7·88	8·00	8·04	8·03	7·91	7·90
Equation (3)	7·70	7·84	7·93	7·92	8·05	8·11	8·14	8·05	8·09
T_m absolute—									
Equation (2)	2585°	2620°	2655°	2645°	2690°	2700°	2700°	2660°	2660°
Equation (3)	2585°	2640°	2660°	2665°	2710°	2730°	2740°	2710°	2715°
Percentage H_2O-dissociation at T_m	2·2	1·8	1·7	1·4	1·4	1·5	1·3	1·2	1·0

Previously, in Part V hereof (*loc. cit.*, pp. 398–400), on the basis of Nernst's figures, the H_2O-dissociation had been provisionally estimated as not exceeding 1·0 per cent.; it would now appear to vary continuously between 2·2 per cent. at the lowest and 1·0 at the highest initial pressure. And, as in the case of the theoretical CO-air explosions, the effect of the new "co-volume" correction for P_m now introduced has been to reduce materially the calculated T_m's in all explosions where the initial pressure exceeded 25 atmospheres, as shown in Table on p. 480.

The progressive increase in maximum temperature with increase of density of the gas mixture is well illustrated by the curve in fig. 2 in which the temperatures plotted are those given by equation (3). It will be noticed that for the two

---

* 'Z. Phys. Chem.,' vol. 87, p. 641 (1914).

Date.	P_i (atmospheres).								
	3.	10.	25.	50.	75.	100.	125.	150.	175.
Calculated T_m in ° C. abs.—	°	°	°	°	°	°	°	°	°
1925	2573	2623	2693	2733	2823	2893	2953	2963	3023
1928	2585	2640	2660	2665	2710	2730	2740	2710	2715

carbonic oxide mixtures there is a rapid rise in the maximum temperature attained as the initial pressure is increased from 3 to 50 atmospheres, after which the further increase is much slower and more uniform. The hydrogen-air mixture shows a similar (but not so rapid) rise in maximum temperature as the initial pressure increases from 3 to 50 atmospheres, which is followed by a period of fairly uniform temperature as the pressure increases to $P_i = 175$.

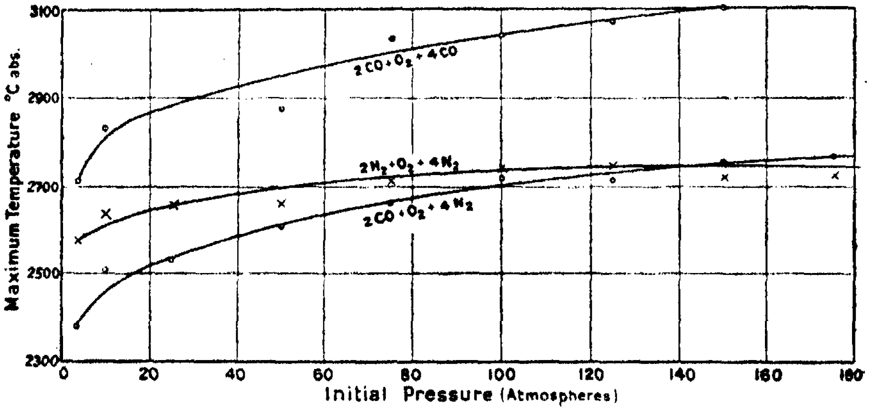


FIG. 2.

In conclusion, I desire to thank Prof. Bone for his constant interest and encouragement throughout my work and for placing all the resources of his laboratory at my disposal for it; also Prof. S. Chapman, of the Imperial College, and Mr. H. T. Tizard for advice in connection with particular matters upon which I have consulted them. Mr. C. M. Smith has also rendered me valuable assistance in connection with the new experimental work involved. Finally, my grateful acknowledgments are due to the Department of Scientific and Industrial Research for a personal grant which has enabled me to devote my whole time to this and cognate researches. The cost of the apparatus used has all been defrayed out of liberal grants placed at Prof. Bone's disposal by the Department of Scientific and Industrial Research and the Government Grant Committee of the Royal Society.

Erosion by Water-hammer.

By Stanley S. Cook, B.A. (Cantab.).

(Communicated by Sir Charles Parsons, F.R.S.—Received January 13, 1928.—
Revised May 23, 1928.)

[Plate 9.]

Corrosion or erosion of the surfaces of metals under water is in many cases found to be actively promoted by their motion through the water, or by the motion of the water over them. In such cases the relative parts played in the action by corrosion and erosion respectively are not readily distinguished. Even where corrosion products are found the action may still be primarily one of erosion, the latter merely creating the conditions favourable to corrosion. It may, for example, have the effect of removing a surface layer or protective scale. On the other hand, where corrosion products are not found the process is frequently claimed to be a corrosive one, the only function allowed to erosion being that of removing the products of corrosion as fast as they are formed.

No doubt in some instances this latter point of view is justifiable, but in many cases the reason for attributing the action to corrosion rather than to erosion is merely that the mechanical nature of the cause is not sufficiently apparent, through a failure to recognise the remarkable ability of water and other nearly incompressible fluids, under certain circumstances, to produce momentary pressures far in excess of the yield-point strength of ordinary materials. The object of this note is to show that in many instances a satisfactory mechanical explanation of such action can be advanced by associating it with the phenomenon known as water-hammer, and, in particular, to show that the same principles afford a satisfactory explanation of the impact pressures produced by drops of water impinging on metallic surfaces.

As a member of the Committee appointed by the Admiralty in 1915 to determine the cause of the erosion of propeller blades, the author, in the early spring of 1917, investigated the hydrodynamic properties of collapsing cavities in an incompressible fluid, and made calculations of the pressures that might arise from the collapsing vortices of cavitating propellers, which being subsequently verified by experimental methods convinced the Committee that the deterioration of propeller blades of cruisers and destroyers by erosion was caused by water-hammer effects resulting from cavitation.* Two years later, in 1919, with the permission of the Admiralty a paper was read by Sir Charles Parsons and the author at the Spring Meeting of the Institution of Naval Architects,† describing the work of the Committee at considerable length, including the above-mentioned calculations and a description of the experimental apparatus by which they were confirmed.

In recent years, with the higher surface speeds and the high vacua generally adopted in steam turbines the leading edges of the moving blades, especially towards the exhaust end of the turbine, are frequently found eroded. Several attempts have been made to account for this erosion. Experimental investigations have pointed to the impingement of water drops on the surface of the blades as being the cause, but the theoretical treatment has not hitherto led to a satisfactory conclusion. The author has recently succeeded in applying methods similar to those adopted by him in 1917 to the calculations of the impact pressures, and has been able to show that such pressures are amply sufficient to account for the observed erosion.

The investigation of the problem of propeller erosion led the author to consider the case when a spherical cavity, in an incompressible fluid, is suddenly permitted to collapse. Such a condition of things might arise in various ways. For example, a propeller blade might cut across a vortex core in such a way as to isolate a cavity and at the same time destroy the vortex motion. Or a sudden arrest of flow of water in a tube by stoppage of the inlet would, by the momentum of the water, tend to break the continuity of the fluid and create cavities which would immediately afterwards tend to collapse. The work done by the surrounding fluid closing in as the cavity collapses is converted into velocity energy, and since, in an incompressible fluid moving symmetrically towards a fixed point, the velocity varies inversely as the square of the distance from that point, this velocity energy will be found mainly concentrated at the reduced surface of the cavity. It was shown that the velocity at the surface $= \sqrt{\frac{2}{3}P(R_0^3 - R^3)/R^3}$, where $P =$ the pressure of the surrounding fluid and R_0 and R the initial and final radii of the cavity. Thus for $R = 1/20\ R_0$ and $P = 1$ atmosphere the velocity at the final surface of the cavity $= 730$ metres per second, and supposing the cavity to collapse finally on a surface with this velocity (we might, for example, consider a hemispherical cavity collapsing on to the surface of a propeller blade in the diametral plane) it was further shown that the water-hammer pressure thus produced would be as high as 10,000 atmospheres or roughly 65 tons per square inch, the pressure finally reached depending upon the ratio of the initial to the final radius of the cavity, and, therefore, upon the original volume of the cavity and the coarseness of grain of the surface upon which it collapses.

---

* Report of the Propeller Sub-Committee (section III) of the Board of Invention and Research—Erosion of Propellers—September 17, 1917.

† "Investigations into the Causes of Corrosion or Erosion of Propellers," 'Transactions of the Institution of Naval Architects' (1919).

In such a case the final act of erosion is due to water-hammer, the concentration merely having the effect of increasing the velocity in the final stage of collapse and making the water-hammer more intense. The pressure due to this water-hammer is determined as follows:—

The pressure generated on an element of surface at its first encounter with water moving at a finite velocity is different from that produced by the steady impact of a moving stream of water at the same velocity. If a moving column of water is suddenly arrested by a fixed surface, without any interposing gas to cushion the blow, there is a sudden arrest of the front layer of the impinging column. The pressure generated by the blow in such a case is only limited by the fact that both solids and fluids possess a slight degree of compressibility, so that the velocity energy of the front layer of the column is converted on impact into potential energy. If V be the velocity of the column of water the kinetic energy of a layer of thickness Δh and unit area of section is $\frac{1}{2}\rho V^2 \Delta h$, where ρ is the density of the fluid. After impact the velocity energy of this layer is destroyed and replaced by potential energy of amount $\frac{1}{2}\beta p^2 \Delta h$ (β being the compressibility). The equivalence of these two amounts of energy gives $p = V\sqrt{\rho/\beta}$ as the pressure instantaneously set up in the front layer of the column. Thus the pressure generated in this simple case of a water-hammer blow is $V\sqrt{\rho/\beta}$. Since for water $\rho = 1$ and $\beta =$ approximately 1/20,000 per atmosphere of pressure, a velocity of only 10 metres per second will, on sudden arrest, give rise in this way to a pressure of 140 atmospheres.

It will be seen that for low velocities the momentary pressure caused in this manner is much greater than the steady pressure due to the momentum of a jet of fluid at the same velocity. The latter is ρV^2, and will only have the same value as the water-hammer pressure when $\rho V^2 = V\sqrt{\rho/\beta}$ or $V = 1/\sqrt{\beta\rho}$, that is to say, when the velocity is equal to the velocity of sound in the fluid in question. For velocities much lower than this, the water-hammer pressure will greatly exceed that due to an impinging jet, for example, at 250 metres per second, or about one-sixth of the velocity of sound in water, the water-hammer pressure will be six times as great as that arising from the momentum of a jet of water steadily impinging at the same velocity.

It should be noted, too, that the water-hammer pressure induced by a suddenly arrested column of water is independent of the length of the column. It follows that drops of water, however minute, will at the first moment of impact produce the same water-hammer effect as large volumes of water, the only difference being in the area of attack and the duration. The duration will depend upon the length of the column, being the time required for the shock wave to travel to the end of the column and back again to the surface of impact. For example, if the column is 1 metre in length the pressure endures for about one seven-hundredth part of a second.

The above considerations show that arising from the mechanical properties of nearly incompressible fluids water-hammer on small areas is a powerful factor contributing to erosion, and that where the conditions are such that cavities may form in the water and subsequently collapse in the neighbourhood of metallic surfaces, there is then the possibility of much higher velocities in the water locally than are normal to the general conditions, and, in addition, the certainty that wherever such high velocities are destroyed by the interposition of a surface, intense momentary pressures will be set up by the water-hammer effect above referred to.

Rapid erosion of the same character as on propeller blades, frequently encountered on the runners, casings and liners of centrifugal pumps, screw pumps and hydraulic turbines, can also be satisfactorily accounted for in this manner.

In the same way the erosion of steam turbine blades may be attributed to the water-hammer of drops of water impinging on the surface of the blades in a high vacuum. On the outer portion of the moving blades at the exhaust end the leading edges are found to be thickly honeycombed with minute indentations of conical shape of varying depths and are sometimes completely perforated, the action occurring on the convex surface of the blade, that is to say, on the side towards which the blade is moving. It occurs chiefly at peripheral speeds in excess of about 120 metres per second, and is of increasing intensity as the speed is increased.

This erosion is attributed to drops of water arising from condensation of the steam by expansion, which drops moving at a lower velocity are overtaken by the rotating blade. It has already been pointed out that the water-hammer pressure due to the impact of a drop is the same as for a larger volume, depending only upon the velocity. The water-hammer pressure calculated as above is 12 tons per square inch for a speed of 120 metres per second and 30 tons per square inch for a speed of 300 metres per second, so that it will be seen to be of the same order as the yield-point strength of materials usually employed.

Many experiments have been carried out with the object of producing such erosion of turbine blades under controlled conditions.* They have, however, employed either high-pressure steam jets impinging directly on blades or bars of metal, or water jets upon revolving test pieces or blades, and are not, therefore, representative of the conditions prevailing in the exhaust-end blading of turbines. Jets were doubtless employed from the point of view that a jet would be more destructive than isolated drops, but the foregoing investigations of the water-hammer effect of drops indicates that a large number of small drops is likely to be the more destructive. Further, there is the possibility of some of the drops being of irregular shape, with the effects referred to below.

In order to investigate the erosion produced by a multitude of minute drops impacting on metal at high velocity, the following experiment has been made : The apparatus consisted of a ring of turbine blades mounted on a shaft and rotated at a high speed within a cylindrical casing. At two diametrically opposed points in one end wall of this casing two sprayers of the "White" oil-fuel burner type were used to project a fine spray of water across the moving blades. The blades were widely separated in the circumferential direction to give the spray access to their surfaces. This was necessary because on account of the high speed of motion of the blades which made the relative direction of impingement of the water drops nearly circumferential. The bladed rotor thus represented an exhaust-end element of a steam turbine from which two out of every three blades were removed. The rotor was 12 inches diameter and the blades 4 inches high. The axes of the water sprayers were situated at a distance of 8 inches from the axis of the rotor.

The apparatus was run for 18½ hours at a speed of 8800 revolutions per minute, so that the speed of impact varied from 140 metres per second at the roots of the blades to 233 metres per second at the tips. By the formula given on p. 483 the corresponding water-hammer pressures are 12·4 tons per square inch to 20·7 tons per square inch.

The blades were of five different materials as given below :—

Material of blades.	Elastic limit. Tons/sq. inch.	Condition after test.
1. "Hecla A.T.V." steel	51	Unattacked.
2. Tungsten steel	40	Rusted and slightly eroded.
3. "Staybrite" steel	16	Considerably eroded.
4. Monel metal	18	Considerably eroded.
5. Mild steel	16	Etched as by sand blast.

---

* 'Brown Boveri Review' (Dec., 1924, April, 1927).

The fine sprays of water playing on the mild steel blades had a peculiar effect. For about ¼ inch across, the blades had the appearance of having been sandblasted, while the remaining parts were rusted. If the erosion be taken as indicated by the raggedness of the blade inlet edge, then the mild steel blades were not eroded as much as the " Staybrite " steel or " Monel " blades.

Fig. 1 (Plate 9) is a photograph of the rotor after test, the blade at the top and the one to the left of it are of " Monel " metal, the next two to the right are of "Staybrite" steel, whilst the next two to the left are of mild steel sheathed with " Hecla " steel on their leading edges.

As has been stated, erosion only occurs on the moving blade. From this it is apparent that it is not due to such drops of water as are carried in the main stream of the steam through the blade passages. Such drops, in fact, issuing from the blades with the same speed as the steam would not be overtaken by the moving blades. It must therefore be attributed to drops of water swept off the fixed blades into the path of the moving blades. Also it is found that erosion occurs most markedly on that part of the moving blade which is immediately opposite to a binding strip of the fixed blades, indicating that drops are blown off the binding strips.

Minute drops of water assume a practically spherical shape on account of surface tension, but drops formed in the manner suggested may be of moderate dimensions, and under these circumstances a slight departure from spherical shape may be expected. With drops of irregular shape, or even if there are minute surfaces of irregularities on the blades themselves, there is the possibility of the isolation of small empty spaces between the drop and the blade at the moment of impact, in which case the behaviour subsequent to impact would be that of a collapsing cavity with the final generation of much higher water-hammer pressures.

Fig. 2 illustrates the progress towards final collapse after impingement, as it might be conjectured to take place with a drop of this character. In the final stages the concentration of the motion towards the centre of collapse at A will produce high velocities at the surface of the entrapped cavity, and finally water-hammer at A.

Further, when a conical depression has once been formed it is clear that it will be rapidly deepened by drops that may strike directly into it or over its mouth, since concentration will then have its full effect in producing a high final velocity.

Another example of what is usually described as corrosion, but is possibly capable in part of a mechanical explanation, is the pitting of condenser tubes.

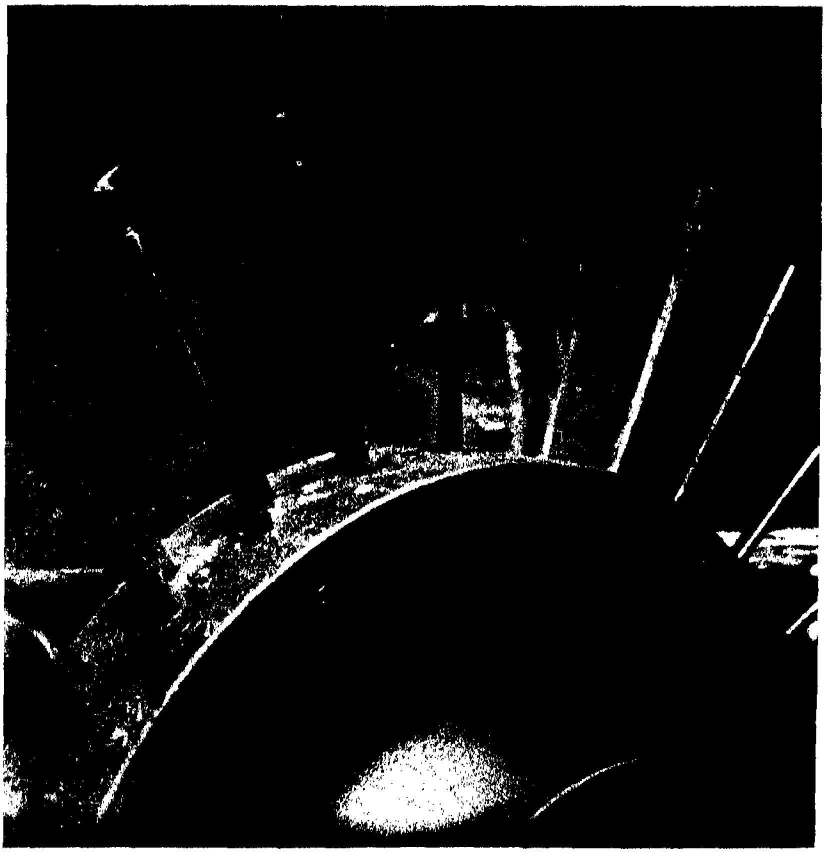


Fig. 1.

In a recent paper by Sir Charles Parsons to the Institution of Naval Architects* some investigations are described in regard to the cause of pitting of condenser tubes, which go to show that irregularity of flow and vortex motion in the water box of a surface condenser may cause considerable variations of flow through the tubes, the flow through a tube being momentarily checked when a vortex in the box moves across its mouth, with consequent breakages of continuity of the water near the inlet end of the tube and subsequent collapse of the cavities so created, producing water-hammer effects. These experiments are at present under continuation.

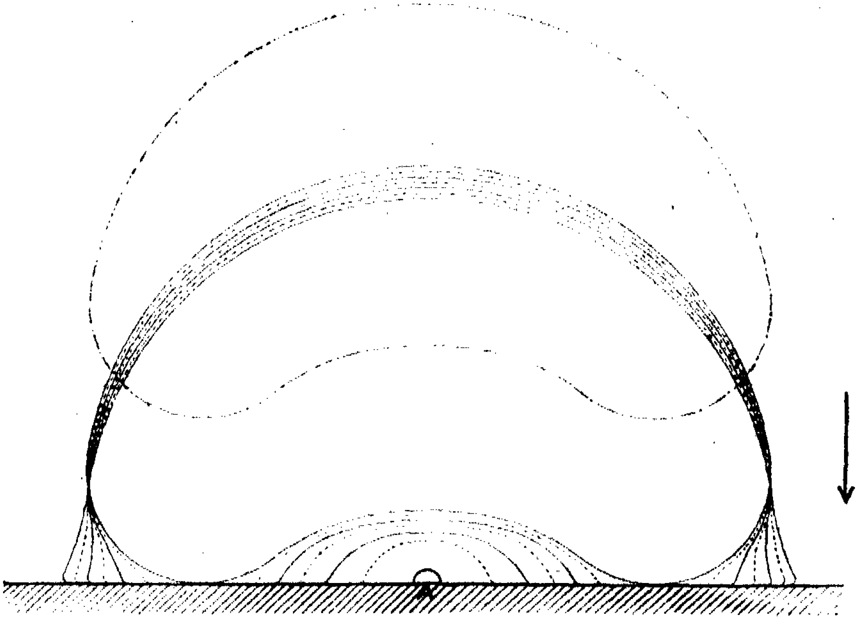


FIG. 2.—Impact of irregularly shaped drop of water against a fixed surface. Successive stages of deformation after impact, showing cavity enclosed at middle of front and rapidly collapsing. The successive contours of this cavity are necessitated by the constancy of volume of the fluid.

Whilst it is generally held that occluded gases play an important part in the pitting of brass condenser tubes, and certainly there are in most cases distinct evidence of corrosion, oxidation and dezincification being found to have occurred near the pits, the causes just mentioned are sufficient to account for pitting of a purely mechanical nature, and it may be that occluded gases compressed to a high pressure and possibly to a high temperature at the final collapse of a cavity are brought to a state of high chemical activity so that both corrosion and erosion result from the same causes.

---

* "Some Investigations into the Cause of Erosion of the Tubes of Surface Condensers" (April, 1927).

The late Lord Rayleigh discussed the case* in which a spherical cavity, instead of being vacuous, contains a small amount of gas, in which case the velocity of the boundary comes again to zero before complete collapse, the whole of the energy of collapse having been converted into the pressure energy of this imprisoned gas. The final volume is shown to be extremely small when the initial pressure of the gas is only a small fraction of that of the surrounding fluid. It was, however, assumed that the compression of this gas would take place isothermally, whereas, in the last stages of the compression at any rate, the interval of time during which it takes place being usually extremely small, compression is probably adiabatic, and a high temperature will be reached as well as a high pressure.

It will be seen from the examples mentioned that the phenomenon of water-hammer and the formation and collapse of cavities in water intensifying the water-hammer in the manner described furnish ample material for the explanation of both corrosion and erosion of metallic surfaces exposed in various ways to such conditions.

---

* 'Phil. Mag.,' vol. 34, p. 94 (1917).

Solid Dipleidoscope Prisms.

By C. V. Boys, F.R.S.

(Received March 22, 1928.—Revised May 1, 1928.)

1.—*Historical.*

In the year 1843 Mr. E. J. Dent wrote a pamphlet on "The Dipleidoscope," a copy of which is in the library of the Royal Society. From this I learned that the instrument known as Dent's Dipleidoscope was invented and patented by J. M. Bloxam, who invented also the gravity escapement. This instrument, manufactured and sold by Dent, had a considerable vogue in the middle of the last century.

As this instrument now is very little known, and as the solid prisms the subject of this paper are a development of Bloxam's invention, it may be well to give in the fewest possible words a description of Bloxam's hollow prism and its use.

Bloxam mounted two plane mirrors silvered at the back behind a plane parallel plate of glass so that the three glasses made angles of 60° with one another, forming in fact a hollow triangular prism (fig. 1).

The silvered mirrors were provided with screw adjustments to eliminate pyramidal error. The exactness of the angles at 60° or their equality is unimportant; their being all parallel to one line is essential. The front plane glass is tinted blue or neutral to bring the two images which are to be observed to equal brightness.

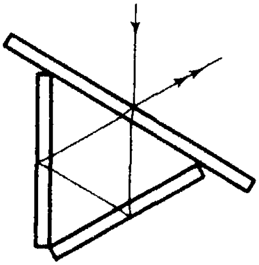


Fig. 1.

When the sun or other luminous object shines on the front or tinted glass from any direction in a plane inclined at 60° to the front glass and parallel to one of the silvered mirrors, part of its light is reflected and part goes through, after which it is reflected from each of the silvered mirrors and part of the remaining light then passes through the tinted glass in the same direction as the light reflected from its own surfaces. If the prism is mounted in an instrument so that its axis is in the meridian and tilted back so as to be parallel with the axis of the earth or thereabouts, and is so positioned that the noon sun is seen in the same direction by either singly or doubly reflected light, then as the sun comes up to and passes this position two suns, a blue and (by contrast) a yellow sun, will be seen moving in opposite directions, and the moments at which the preceding limbs, and afterwards the following limbs, make apparent contact can be observed with great precision. If the telescope from a sextant is used in addition the observation can be made with an accuracy of a fraction of a second and one of the paler dark glasses of the sextant may be used to bring the light down to an intensity suitable for observation.

The length of the prism is such that an object within the range of the ecliptic is easily visible, but it is just possible to observe the transit of a star so far away as Capella.

This instrument was so neat and convenient that it has always been a wonder to me that it was not more widely used. At the present time from a purely utilitarian point of view wireless time signals would obviously be preferred. There are two examples in the Science Museum, South Kensington.

2.—*The Solid Prism.*

Fig. 2 shows a solid prism of 60° with an incident and emergent ray, which latter is due either to single or double reflection, and it might appear that the solid prism could equally well be used in a dipleidoscope, and if ever free from pyramidal error, would always be so. I have adjusted so many of the Bloxam prisms to correct them for pyramidal error that it is clear, no doubt from interference by users, this error was occasionally found. An instrument which I have had for more than 50 years has, however, remained free from this.

The solid prism of fig. 2 would be perfect with monochromatic light, but as with white light the emergent light is a spectrum there is no definite doubly reflected object to be seen.

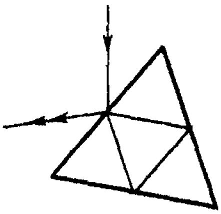


FIG. 2.

So far as I know the general problem of the solid dipleidoscope prism has not been considered, and it is the object of this paper to develop the subject and to show how, incidentally, a very convenient and accurate method of testing for pyramidal error in prisms generally is a corollary, and to produce a domestic transit instrument more convenient in use than Bloxam's, which may be worth describing, even though it could have no general use now on account of the universally available wireless signals. I also show that the solid dipleidoscope prism might be used with advantage to increase the precision of our best astronomical instruments.

I have good reason for supposing that the method of testing for pyramidal error is not generally known, for it is not mentioned in the article on the subject in Glazebrook's 'Dictionary of Physics,' and when in 1924 Messrs. Hilger made for me the two prisms which I am about to describe, they did not know of the test, and I was able to show to them at once that neither prism was quite truly prismatic. They corrected one, and I was glad to retain the other uncorrected to show the delicacy and simplicity of the test.

In order that a solid dipleidoscope prism may not produce a spectrum the light on emergence after an even number of internal reflections must be bent in the opposite direction to that in which it was bent on entry, and to the same extent. Fig. 3 shows in thick lines a prism ABC, with a pair of parallel incident rays *p*, *p*, and an emergent ray *e* common to both. The right-hand incident ray enters the prism being bent to the right, then after two internal reflections it is bent equally to the left and follows the path *e*. The course of the internal rays and the condition for equal refraction is best followed in the figure by the use of the two ghost prisms A′BC and A′B′C, these being reflections of the real prism about the two reflecting faces BC and A′C. In order that the emergent ray *e*′ may be free from colour the emergent ghost face A′B′ must be parallel to AC. Now the triangle A′B′C is merely the triangle ABC turned through twice the angle at C, and such turning has brought the face AB parallel to the face AC. Hence the angle A is double the angle C, and this is the general condition for a solid dipleidoscope prism.

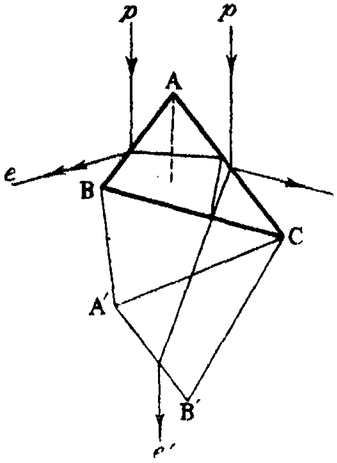


FIG. 3.

The angle A may vary between fairly wide limits, but as the entering beam must be of the same width as the emergent beam, and the face AB is for all useful prisms shorter than the face AC, it is best to utilise the whole of the face AB, and then the incident beam must occupy a portion of AC equal in length to AB.

Two prisms are of special interest, the first in which the portion of the face AC occupied is bounded by the vertex A and the second when it is bounded by the far angle C. The prism shown in fig. 3 is intermediate between these two and it also has a peculiarity to be referred to later. If the angle A is made more obtuse than that of the second prism, only a portion of the face AB is utilised. If the angle A is made a right angle and the other two are necessarily each of 45°, then the common reflecting prism may be seen to have the dipleidoscope property, but it is a very bad example because only a very narrow beam is available and the intensity of the singly reflected ray is much greater than that of the ray which has passed through the prism. This prism, however, admits of observation from both sides at the same time.

3.—*The Polarising Prism.*

The first prism is illustrated in fig. 4. In this the incident light strikes the prism faces at the polarising angle. The partially polarised ray entering the prism is perpendicular to the base from which it is reflected along the same path. It is then reflected at the point at which it entered the prism where on reflection it becomes completely polarised. It then leaves the prism by the near face on the same path as the direct reflected ray from that face. Both rays are completely polarised in the plane of incidence. If i is the angle of incidence and r that of refraction it will be seen that all the angles as given in the figure are necessary consequences of the combined conditions,—complete polarisation and solid dipleidoscope prism. The total path traversed by any ray within the prism is equal to twice the distance from the vertex to the base.

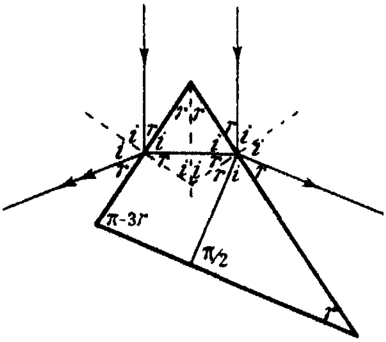


Fig. 4.

The calculation of the intensity of the two beams of emergent light is simpler in this case than in any other, because there is light polarised only in the plane of incidence to consider. Combining Fresnel's sine formula for intensity $\sin^2(r-i)/\sin^2(r+i)$ with the relation $\tan i = \mu$, and writing down the result it appears that the intensity of the light reflected externally in relation to half the incident light is $\{(\mu^2-1)/(\mu^2+1)\}^2$ which with $\mu = 3/2$* is $(5/13)^2$. The light entering the prism is $(2\mu/\mu^2+1)^2 = (12/13)^2$. The fraction of this that is reflected perpendicularly from the base is $\{(\mu-1)/(\mu+1)\}^2 = 1/25$. The fraction of what is left reflected internally from the far face is $\{(\mu^2-1)/(\mu^2+1)\}^2 = (5/13)^2$, and the further reduction of intensity on emerging at the front face is $(2\mu/\mu^2+1)^2 = (12/13)^2$. The intensity then of the emergent light is $\left(\frac{2\mu}{\mu^2+1}\right)^4\left(\frac{\mu-1}{\mu+1}\right)^2\left(\frac{\mu^2-1}{\mu^2+1}\right)^2$, that is, it is equal to the externally reflected light × $\left(\frac{2\mu}{\mu^2+1}\right)^4\left(\frac{\mu-1}{\mu+1}\right)^2$. The first of these factors is $(12/13)^4 = 0{\cdot}72$ and the second is $0{\cdot}04$. This would make the transmitted light only 3 per cent. of the externally reflected light which would be very unsuitable for good observation. I have, however, overcome this difficulty by silvering the base. I have thus raised the intensity of the transmitted light until it is almost three-quarters of that of the externally reflected light and the two beams of light are sufficiently nearly equal for good observation. The fact that the two emergent beams are completely polarised in the same plane may be of the greatest advantage as will appear in the seventh section of this paper.

---

* The actual value of μ for the two prisms made for me by Messrs. Hilger is stated by them to be 1·51230 for a yellow ray and the angles are correct for this value. I have used the value 3/2 in the discussion for simplicity and the figures are correct for this value.

It may be worth while to mention that when the prism is so placed as to give the coincidence seen in the front or short face, light which strikes the front face near the vertex and enters the prism emerges after six internal reflections at the back face in the same direction as the light reflected from that face and without colour, and so an observer looking at this face can witness a transit of the sun at the same time as another observer looking into the short face. In fig. 5 I have shown the prism with the six ghost prisms from which it is apparent that as A'''C' is parallel to AB the light must emerge without colour after following the track indicated in the actual prism. As three of the reflections are from the silvered base and one from AC is total the serious loss of light is from two reflections only and the image is easily seen.

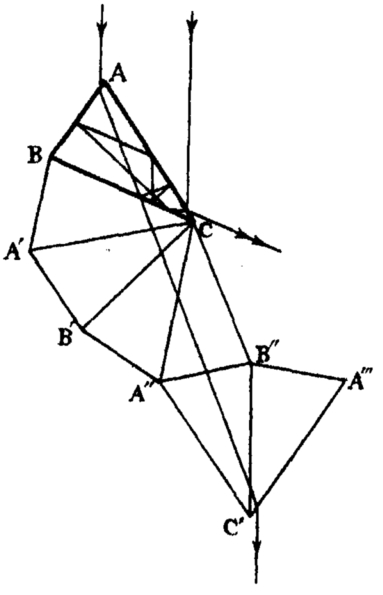


Fig. 5.

A third coincidence of two spectra can be seen at the same time by a third person in the long face by rays which have followed two different internal tracks, one with an even and one with an uneven number of reflections. The two spectra face the same way. It does not seem worth while to trace this out.

If the polarising prism is to be used for front face observation only, then the acute toe of the prism is redundant, and all beyond a plane perpendicular to the base and cutting the long face so as to leave an amount equal to the short face may be absent.

4.—*The Squat Prism.*

If the vertical angle of a polarising prism is gradually enlarged and the far angle is maintained at half the vertical angle, then the beam of light entering the long face and emerging at the short gradually occupies lower positions, passing through that illustrated in fig. 3 until at last it comes down to the far angle. Any increase after this in the vertical angle merely wastes part of the glass and a narrow beam is the result. When it is terminated by the far angle the best use is made of the glass, the widest beam of light is obtained, and as will be seen fair equality between the intensities of the two beams is obtained without silvering the base. The squat prism having these properties is the most suitable unless the special features of the polarising prism are required. The squat prism is illustrated in fig. 6. If the entering beam on the long face which emerges from the short reaches only just as far as the far angle, this is another way of saying that the second reflected ray is parallel to the base. Then calling i the angle of incidence and θ the complement of i, r the angle of refraction and ϕ the complement of r, all the angles of the figure can at once be written in as illustrated. From this it will be seen that the first internally reflected ray is parallel to the primitive ray, and the total distance traversed by any ray within the prism is equal to the length of the base.

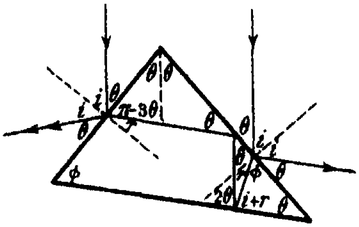


FIG. 6.

Since
$$\phi + 3\theta = \pi = 3i - r,$$
$$\cos 3\theta = -\cos\phi = -\cos\theta/\mu.$$

Hence,
$$\cos^2\theta = \sin^2 i = (3\mu - 1)/4\mu = \mu^2 \sin^2 r.$$
$$\sin^2\theta = \cos^2 i = (\mu + 1)/4\mu.$$
$$\cos^2\phi = \sin^2 r = (3\mu - 1)/4\mu^3.$$
$$\sin^2\phi = \cos^2 r = (2\mu - 1)^2(\mu + 1)/4\mu^3.$$

Considering first light polarised in the plane of incidence and substituting in Fresnel's sine formula one or other of the above-mentioned values, it appears that the intensity of the light externally reflected from the two faces of the prism is equal to that of the incident light multiplied by $(\mu - 1/\mu)^2 = 1/9$, if $\mu = 3/2$, while that transmitted is $(2\mu - 1)/\mu^2 = 8/9$.

At the internal reflection from the base of the prism, the coefficient of reflection is, if $\mu = 3/2$, equal to 0·04356. The other internal reflection is total. On multiplying the three factors 8/9, 0·04356 and 8/9 together the intensity of the emergent beam is found to be 3·449 per cent. of the incident light polarised in the same plane.

Similarly the factors for light polarised across the plane of incidence deduced from Fresnel's tangent formula for intensity $\tan^2 (r - i)/\tan^2 (r + i)$ are for external reflection

$$\left(\frac{\mu - 1}{\mu}\right)^2 \left(\frac{\mu(\mu - 1)}{\mu^2 + 2\mu - 1}\right)^2 = \frac{(\mu - 1)^4}{(\mu^2 + 2\mu - 1)^2} = \left(\frac{1}{17}\right)^2 = \frac{1}{289} = 0\cdot346 \text{ per cent.},$$

while that transmitted when entering and again when leaving the prism is $4\mu^2 (2\mu - 1)/(\mu^2 + 2\mu - 1)^2 = 288/289$.

At the internal reflection from the base of the prism the coefficient of reflection is, if $\mu = 3/2$, equal to 0·0366. On multiplying the three factors 288/289, 0·0366 and 288/289 together the intensity of the emergent beam is found to be 0·0363. Collecting then the two components of the two beams of light, one externally reflected and the other emerging after two internal reflections, the quantities are :—

Polarised.	Externally reflected.	Transmitted.
In plane	0·1111	0·0345
Across plane	0·00346	0·0363
	0·1145	0·0708
Fraction of incident beam........	0·0572	0·0354

thus the externally reflected light is brighter than the transmitted light, but not to such an extent as to interfere with accurate observation. This is quite obvious when the prism is used. If, however, the pair of images are viewed through a Nicol prism then, as this is turned from the position where it allows light polarised in the plane of incidence to pass to that in which light polarised across this plane can pass, it is also very evident that the externally reflected light instead of being far the brightest becomes far less bright than the transmitted light. It is easy to make a good photometric balance at an intermediate position. I shall return to this question in the sixth section of this paper.

The squat prism has the curious property that, if placed for coincidence under observation at the AB face, the rays entering the face AC near the angle A and reflected internally from the face BC and totally from AB emerge from AC as a spectrum on which the direct reflected light from that face is superposed, so that the colour for which the prism is correct is directly evident. The series of angles marked in fig. 7 can all be put in by inspection. As the light is bent as much on emergence as it is on entering the prism this is in effect placed in the position of minimum deviation and the equivalent angle of the prism is $3\theta - \phi$.

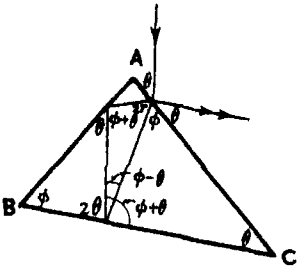


FIG. 7.

The squat prism does not need to have the base silvered as the reflection from AC is total, whereas in the polarising prism this latter reflection cannot be total and the silvering of the base becomes necessary to give fair equality between the brightness of the two images. The prism shown in fig. 3 is intermediate between these two, and has such an angle A = 74° that the transmitted beam is totally reflected when seen to the right of the position of coincidence and only partially reflected when seen to the left, and such result is in general undesirable. Euclid's triangle with A = 72° is dipleidoscopic with light from two directions.

5.—*Testing Pyramidal Error.*

It will be evident that if any dipleidoscope prism is not truly prismatic, that is, if the three faces are not truly parallel to one line, but would instead if produced ultimately meet in a point, then the two apparent suns or stars will not pass exactly across one another, but one will pass over the other, as happens also with a sextant out of adjustment.

As it is unimportant for the purpose of this test that the transmitted light should be free from colour there is no necessity to confine the method to dipleidoscope prisms. The squat prism, as already mentioned, gives a superposition of a colourless externally reflected image upon a transmitted spectrum, and this is quite suitable for observation of pyramidal error as well as for the observation of the colour for which the angles are correct. This indicates that prisms other than dipleidoscope prisms may be tested in the same way, and if of more than three sides the superposed images should be due to reflections or refractions at three faces at a time only. Any other faces can be rendered sufficiently non-reflecting by placing on them black paper wetted with water or a more refractive liquid.

6.—*Testing Fresnel's Equations.*

I have already indicated that the solid dipleidoscope prism may be used for comparing the intensity of the illumination of the two images either with the light unanalysed or with the light polarised in one or other of the principal planes. It so happens that these prisms afford the most perfect possible means for comparison, for the two images of an illuminated rectangular opening can be brought edge to edge, and as one is a reversed image the conditions for exact photometric comparison of equality are perfectly realised. Reference to the late Lord Rayleigh's well-known investigation* of the exactness with which Fresnel's formulæ are justified by experiment will show what a difficult experiment it was that he made. The solid dipleidoscope prism at once removes all difficulty, and it may be used to examine the effect of a condensed layer of contamination on the glass which impairs the sharp discontinuity at the surface. On the other hand, only certain angles can be tested.

Whether a Nicol prism is used for viewing or not equality can be obtained by the use of a revolving sector in the brighter of the two beams before they meet the prism. Then when the setting for equality is found for the prism in any particular state of contamination, the reflecting and transmitting faces can be cleaned chemically or by polishing or be contaminated one at a time and the effect on the balance be determined. For instance, the effect of contamination on the face AB would be to reduce the direct reflected light and increase the transmitted light, thus destroying the balance. Similarly, the other faces could be cleaned or contaminated and the effect be observed. In the case of the polarising prism the fact that observations are necessarily made at the polarising angle exactly has the double advantage that the necessity for accurate angular settings is avoided, and the degree of completeness of polarisation at this angle is so easily observed. When using the instrument containing the polarising prism most striking colour changes from orange to blue take place in the transmitted beam as it passes the position of coincidence as will be described. Meanwhile beyond making a few preliminary tests which have satisfied me that these prisms are peculiarly well fitted for this investigation, I have not as yet made systematic tests but hope to do so. The degree of polarisation of the light of the sky may similarly be examined.

---

* 'Collected Works,' vol. 2, p. 522.

7.—*The Window Transit Instrument.*

Going back to the Dent Dipleidoscope illustrated in section in fig. 1, it will be seen that the observer must be in a position with his back somewhat sideways to the sun. It cannot therefore be placed in the window of a room with a more or less southern aspect or in a room at all; the observer must go outside to use it. It would be more convenient for domestic observation if it could be placed in the window and be observed from inside. The solid dipleidoscope prism makes this possible, and as I have made the instrument two other inconveniences of the Dent instrument are avoided. It is not necessary to look up at an uncomfortable angle when observing the mid-summer sun or a star with a higher northern declination, but, as will be seen, the observer may sit down and look downwards at a convenient angle. With the Dent dipleidoscope even though a small fraction only of the intensity of direct sunlight is available, yet what there is is too bright unless cut down with a dark glass, such as one of the shades of a sextant, and when the sun is variously obscured by thin cloud the suitable depth of shade is constantly altering. As will be seen in my instrument no dark shade is necessary, and the fraction of the light observed can be varied continuously over the whole necessary range.

Fig. 8 is a plan of the instrument taken from a direction perpendicular to the sun's or stars' elevation, and fig. 9 is an elevation partly in section as seen from a north-west direction. A is a heavy brass casting with trunnions BB which are supported in V's in an east-west direction and level as in a transit instrument. C is a polarising dipleidoscope prism resting on the flat ends of three supporting screws 1, 2, 3, and against the ends of two other screws 4 and 5 and with its acute angle resting also against the pillar 6. The last three contacts are as near the lower surface of the prism as is practicable. The prism is pressed against all six contacts simultaneously by the spring piece D carried on a pillar and with its free end bent down and pressing against a small inclined facet ground on the top corner of the prism. The tilt of the casting A can be adjusted by a screw E bearing on a block below so that the front edge of the prism is square to the sun. Both beams of light leaving the prism are parallel to the surface of the casting A, and are received upon the inclined mirror F made of dark glass from which a fraction only is reflected into the sextant telescope G magnifying three times. This is mounted so that its inclination may be varied as is made sufficiently clear by the drawing without further explanation, but the telescope arm carries an eccentric pin H against which the dark glass F rests. Its lower end rests upon the end of a screw I, which is as far from the axis of tilt of the telescope as is the pin H, and so when the telescope is tilted more or less the mirror F moves through half the angle. Euclid III, 20. Thus the two images of the sun are seen centrally in the field of view of the telescope however it is tilted. The telescope and mirror are shown in the figure at the polarising angle, but the plane of the last reflection is perpendicular to that of the paths of the rays in the prism. If the polarisation were quite perfect the sun would disappear just at the meeting point of the two images. While never invisible in this way the images are so obscure, even when the sky is perfectly clear, that the sun's images are clear but without the slightest dazzle. If the sky is not so clear a slight tilt will make the sun appear bright enough.

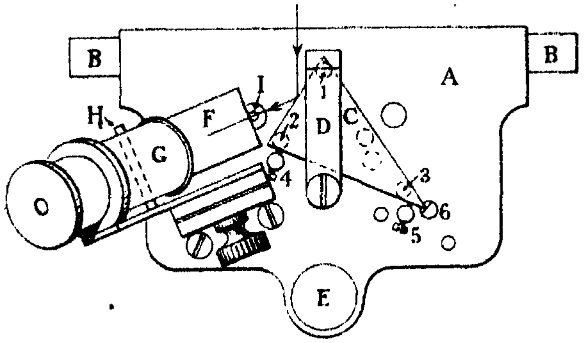


FIG. 8.

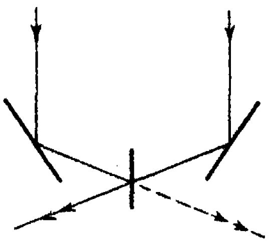


FIG. 10.

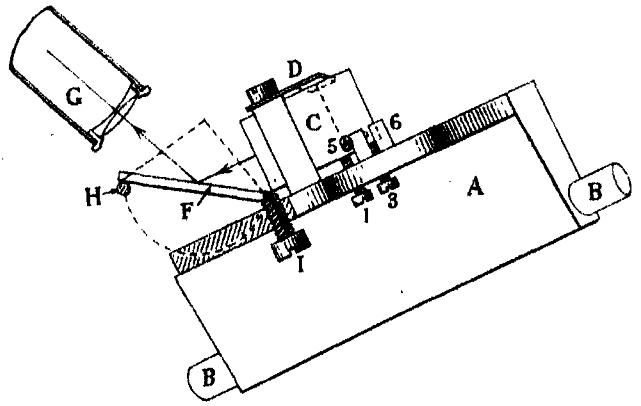


FIG. 9.

It will be evident from the description of the instrument that the screws 1 to 5 give all the desired adjustments to the prism. The screws 4 and 5 are used to bring its plane of coincidence truly perpendicular to the trunnion axis. This is done by viewing a distant mark before and after reversal of position in the V's exactly as with a transit telescope. If a star is to be observed the dark mirror F can be replaced by an ordinary reflecting prism shown dotted in the figure. There is no loss of light then due to cross polarisation.

As it might be desired to set up the instrument in a window with a south-east aspect I have made all the screw holes and conical pillar holes in duplicate in the reversed positions as shown. The prism then can be held in a reversed position and the telescope now on the western side of the prism is placed conveniently for observation. The pillars concerned in the positioning contacts 4, 5 and 6 are all made very slightly conical and they fit into corresponding holes reamed conical in the brass casting. A slight tap with a hammer removes each one and a second tap fixes it in the opposite hole securely. This practice common in the gunsmith's art is not used in scientific instruments so much as it might be. It may be worth while to mention that I have obtained a beautiful smooth motion for the telescope by facing one of the discs, which are drawn together with light pressure, with a layer of cork, such as is used for the tips of cigarettes. For these I have to thank Messrs. Alexander and Alexander, who cut cork into sheets of incredible delicacy and perfection. They kindly gave me a quantity which I hoped would enable me to overcome a difficulty in another instrument. While not successful for that purpose it gives an exquisite smooth and light friction and is most suitable when used as described.

The externally reflected and internally transmitted images of the sun are very different in colour but appear quite equal in intensity of light. The second of the two, which appears to travel in the same direction as the sun, comes into the field of view of an ill-defined orange colour. When it has approached the other image to within about a diameter it has changed to a bluish colour. Then as the two images begin to cross the blue deepens. When the two overlapping images have passed their central position the blue becomes still clearer and the other appears a sort of orange and the overlapping part seems to be white, which was not the case before the central position. When the images have completely separated the blue sun becomes a still more beautiful and clear blue like the blue of the chicory flower, while the externally reflected sun is still a poor orange.

The observation for time is usefully accurate. Watching the two images of the sun pass within sound of a regulator clock there is no question as to the particular tick of the clock at each contact, the half second may be imagined. With more telescopic power some increase in accuracy might be attained.

It may be of sufficient interest to mention that in the year 1875 or 1876 I invented a dipleidoscope in order to obtain a better observation from inside a room than a mere shadow of a vertical edge. I showed this to an astronomical friend who told me of Dent's dipleidoscope of which I had not heard, and he gave me the one to which I have referred. As Dent's instrument was certainly the neater of the two, I said no more about mine. Fig. 10 without further explanation shows how it worked. The central glass was tinted to bring the two images to equality.

As is shown by the dotted line a second observer can witness the same transit at the same time.

8.—*Improvement of Astronomical Instruments.*

The astronomical instrument has in the course of generations been brought to so high a state of perfection that astronomical observers may well look sceptically on any suggestion for replacing the most precise of all—the transit instrument—by any new kind of construction. I believe, however, that more precise observations of time and declination could be made by a new type of instrument with two solid dipleidoscope prisms. Before showing what I propose it may be well to consider from the designer's point of view what is the problem of the big transit instrument. To fix ideas imagine a transit instrument with a focal length of 8 feet and an aperture of 8 inches. One second of arc is represented by a distance at the focus of one two-thousandth of an inch and this corresponds to one-fifteenth of a second of time for an equatorial star. The object glass is held in a brass cell which expands three times as much as this for every 10° C. that it changes in temperature. The object glass is secured in its cell so as to eliminate the effect of expansion with the best skill and experience of the instrument maker, and even if the rest of the structure were perfect the question would still remain how certainly can the user of a big transit instrument be sure that, as it changes in elevation, the relative positions of the optical centre of the object glass and the grid of cross wires cannot have changed by a small amount of this order since the last observation with the mercury horizon, or since the last verification with the collimators. I am not in a position to give an answer, but I hazard the suggestion that there may be accidents of position which may account, for instance, for the slight irregularities of transit observations of time when tested against the Shortt clock. By this I mean that observations of the actual transit over the wires may be more accurate than the deduced transit across the meridian, and the ambiguity may be simply due to want of exactness in the relative positions of parts of the instrument.

From the description of the window transit which I have made it will appear that the positioning of the solid dipleidoscope prism is a simple matter. If the optical glassmaker and the optician between them can produce one of these prisms of the necessary size which is optically perfect, or as perfect as any other optical work, then observation with such a prism and a 4-inch viewing telescope would be equal to a transit instrument observation with an 8-inch object glass› and there would be no double 4-foot cantilever holding a heavy object glass at one end and the grid of wires at the other. The telescope would be a mere viewing telescope. For observations on objects having sensible discs the ordinary eye-piece would be suitable and two contacts would be observed. For stars, as it is very difficult to observe the exact moment when the two apparent stars cross one another, a slightly cylindrical lens in the eye-piece should be used to make the star appear as a short line.

Similarly for altitude observations either direct or from an artificial horizon a second prism with its viewing telescope independently carried would be used. Neither prism would have any possible ambiguity of position due to differential expansion or to sag. This is not the time or place to go more fully into questions of design, but a tubular trunnion axis carrying a divided circle in its centre would be supported as in the case of the usual transit instrument. The trunnions would each carry a prism in an overhung rigid mounting with the necessary adjustments. The two telescopes each balanced, of course, would be carried on bearings on the pillars of the instrument or on independent pillars. Alternatively the trunnion axis might contain the prisms within or between the actual trunnions.

Two objections will naturally be raised to such an instrument: (1) The loss of light would make it unsuitable for observations on faint stars, but the loss of light would be a very great advantage where stars down to the fourth or fifth magnitude are being observed at night; (2) there can be no sequence of observations as on a grid of cross wires. This again is not altogether a disadvantage because with a good observer the very consistency of his results give the appearance of a fictitious accuracy. If the object glass is subject to any uncertainty of position relative to the cross wires, the observer is merely repeating his observations of the consequent error without knowing it.

Similarly a solid dipleidoscope prism could be used in a floating zenith instrument with two fixed viewing telescopes one north and one south.

My suggestion is that if the glass maker and optician can produce the prisms, then the greatest possible accuracy in astronomical observations will be attainable with these prisms, and the telescope will take the second place as a mere viewing machine.

I fully recognise the difficulty of producing the necessary prisms but not the impossibility, because it depends upon how much a thing is wanted whether difficulty is called impossibility or is overcome. I do not think it well, therefore, to lengthen this paper by discussing these points now. I would only say that prisms to be used with a 4-inch lens should take in the whole 4-inch beam in the direction of measurement, but that in a transverse direction they need not be of the full thickness. A thickness of 2 inches would be sufficient and the effect of the smaller thickness would be to increase the possible definition to a small extent, and to hurry very greatly the acquisition by the prisms of the surrounding temperature and diminish the internal strains and consequent double reflection.

9.—*The Prismatic Astrolabe.*

The observation of contacts of direct and reversed images as they meet and pass one another, already well known in the case of the sextant and as used with the Bloxam prism, is made use of also in Claude and Driancourt's prismatic astrolabe. This well-known instrument, which has proved so successful especially for surveying purposes, is extremely limited in its scope. By its aid stars may be observed only as they cross the 60° altitude level, but this altitude of 60° can be determined with great accuracy without dependence on divided circles, levels or accurate settings, it being a consequence of the angle of the prism which is invariable. A mercury artificial horizon is an essential adjunct.

The prismatic astrolabe suffers from one defect and that is the use of the two halves of the object glass for the formation of the two separate images, so if the optical work is perfect the limit of definition cannot be better than half that due to an object glass of the aperture used. This reduced definition is, however, exactly balanced by the doubled relative motion.

The solid dipleidoscope prism as described cannot be used with a prismatic astrolabe with a view to obtaining the full defining power of the object glass as the mercury horizon has already reversed the vertical component of the apparent motion of the star, but a corresponding prism with a single internal reflection would be suitable.

I have shown in fig. 11 how such a prism, telescope and horizon could be arranged to be used in the same way as the prismatic astrolabe. I have chosen a particular angle of altitude for the star of 35° which leads to the construction of a suitable prism, but there is a possible range from below this up to about 45°, beyond which the glass is so wastefully used as to be inconvenient. I have put in all the material angles which follow from the arbitrary selection of

35° for the altitude α of the star. From these it will be seen that the prism is isosceles while the small angles at the base are equal to the altitude α of the star. That the front face is vertical and the top face slopes down at an angle equal to $2 \times (45° - \alpha)$. Also that the telescope looks down at an angle equal to $45° + 3\alpha$. The very small fraction of the glass of the prism not utilised is shown shaded.

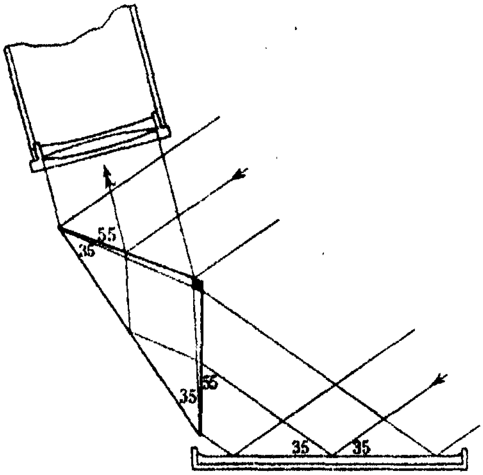


FIG. 11.

The calculated intensity of the light reflected from the top face of the prism is 4·30 per cent. of the incident light, while that reflected from a mercury horizon and passing through the prism is 91·6 per cent., giving such a disparity in brightness as to be most unsuitable for accurate observation. The prism would have to be made of heavily clouded glass to bring the two images to approximate equality.

A better plan is to use olive oil instead of mercury for the horizon. The calculated intensity of the oil-reflected and transmitted beam is then 5·49 per cent. or sufficiently near to 4·30 per cent. for good observation. A very slight tinting of the glass of the prism would bring the two to equality if it were worth while. It so happens that the angle selected gives almost complete polarisation in the light reflected from the oil.

A prismatic astrolabe made with a prism of about the angle indicated and an oil horizon would have the advantages that the definition due to the whole aperture would be attainable, the oil is less sensitive to disturbance than mercury and there would be a larger selection of stars. Claude and Driancourt lay stress on the value of the comfort obtained by looking into a level telescope as against one pointing upwards, such comfort being an essential factor for accurate observation. Looking downwards at an angle of 75° is even more convenient than looking horizontally, and microscopes as ordinarily placed bear witness to this. As a set off to these advantages there is a great loss of light so that brighter stars would have to be selected or observation made when the twilight was less. Also the correction for refraction would be two and a half times as great.

The Mechanism of Chemical Change. I.—Promotion and Arrest of the Mutarotation of Tetra-Acetylglucose in Ethyl Acetate.

By T. Martin Lowry, F.R.S., and G. Glyn Owen.

(Received March 28, 1928.)

In the whole range of organic reactions there is no group which is better adapted for a study of the mechanism of chemical change than that in which changes of structure are accompanied by changes of optical rotatory power, giving rise to the phenomenon which is now generally known as "mutarotation."* Thus, apart altogether from the special interest which attaches to the facile changes which give rise to this phenomenon, the experimental conditions are particularly well adapted for making critical observations, in the first place, because the progress of the change can be followed with ease and accuracy by polarimetric observations, and in the second place, because the conditions under which the change takes place are exceptionally favourable for work under "clean" conditions. Thus, the solutions under examination need not be contaminated by contact with electrodes, or other measuring instruments, or even by the introduction of pipettes, but can be left enclosed in a polarimeter-tube at constant temperature from the beginning to the end of the experiment. Up to the present, experiments of this character have been made mainly with two groups of compounds, namely, (i) the reducing sugars, (ii) nitro-derivatives and α-diketonic derivatives of camphor; but wide extensions in other directions may be looked for in the future.

The mutarotation of freshly prepared solutions of the reducing sugars has been known for about 80 years,† but it is only recently that the underlying chemical changes have been the subject of detailed investigation. Dubrunfaut's experiments (and all other experiments on mutarotation during a period of about half a century afterwards) were limited to aqueous solutions. When, therefore, physical explanations (such as the progressive decomposition into "chemical molecules" of complex hypothetical "crystal-molecules") were found to be untenable, an obvious chemical explanation of the changes of rotatory power was forthcoming in a progressive hydration of the anhydrous sugars, *e.g.*, $C_6H_{12}O_6 + H_2O = C_6H_{14}O_7$, since the analogous lactones also exhibited mutarotation when hydrolysed to the corresponding acids according to the equation $C_{10}H_{10}O_6 + H_2O = C_6H_{12}O_7$.*

---

* Lowry, 'Trans. Chem. Soc.,' vol. 75, p. 213 (1899).

† Dubrunfaut, 'C. R.,' vol. 23, p. 38 (1846).

The more modern view, that the mutarotation of the sugars is due to a reversible isomeric change, emerged only gradually. The first important advance was made when Tanret,† by evaporating an aqueous solution of glucose at 100°, isolated a form of the sugar which reached the equilibrium-rotation of + 52·5° by an *increase* of rotatory power from + 22·5°, instead of by a *decrease* from + 105·2°; but it was assumed by Tanret‡ himself, as well as by Lippmann,§ that the three rotatory powers + 105·2°, + 52·5° and + 22·5° represented three isomeric forms of the sugar, of which the first and last were converted completely (and therefore irreversibly) into the sugar of intermediate rotatory power. It was therefore not until the mutarotation of nitrocamphor was discovered in 1899 that the theory of *reversible isomeric change*, or "dynamic isomerism," which had been worked out experimentally by Butlerow in 1877,|| was applied to the mutarotation of the sugars. This application was based upon the observation¶ that freshly prepared solutions of nitrocamphor exhibited analogous changes of rotatory power in a series of 15 solvents, from which water was excluded, however, because the nitro-compound was not soluble in it. Since this new mutarotation normally took place in the absence of water, it was clear that it could not be due to hydration, but must be attributed to some chemical change in which the solvent was not directly involved, *e.g.*, to a reversible isomeric (or polymeric) change. It was therefore suggested that the changes of rotatory power (which were always from left towards right) were due to a change of structure similar to that which

---

* E. Fischer, 'Ber. D. Chem. Ges.,' vol. 23, p. 2626 (1890).

† 'C. R.,' vol. 120, p. 1060 (1895).

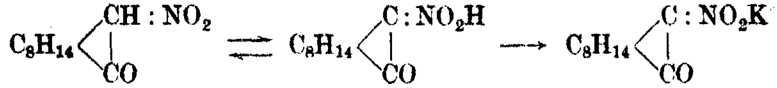
‡ 'Bull. Soc. Chim.,' vol. 15, p. 195 (1896).

§ 'Ber. D. Chem. Ges.,' vol. 29, p. 203 (1896).

|| 'Ann. d. Chemie,' vol. 189, p. 44 (1877).

¶ 'Trans. Chem. Soc.,' vol. 75, p. 211 (1899).

occurs when the lævorotatory nitro-compound is converted into its dextrorotatory salts



$$C_8H_{14}\left<\begin{matrix}CH:NO_2\\ |\\ CO\end{matrix}\right. \rightleftharpoons C_8H_{14}\left<\begin{matrix}C:NO_2H\\ |\\ CO\end{matrix}\right. \longrightarrow C_8H_{14}\left<\begin{matrix}C:NO_2K\\ |\\ CO\end{matrix}\right.$$

Normal nitrocamphor. *Pseudo*-nitrocamphor. Potassium salt.

A series of anomalous observations on solutions in chloroform showed that, although the solvent is not directly involved in this change of structure, mutarotation in inert solvents is not spontaneous, since solutions in chloroform often exhibited an "arrest" of mutarotation extending over a period of several days or weeks.* When, therefore, Purdie and Irvine,† by preparing a sugar which was soluble in organic solvents, obtained direct confirmation of the view (which had been justified previously by analogy with the mutarotation of nitrocamphor) that water does not play any stoichiometric part in the mutarotation of the sugars, it was still possible to contend‡ that the mutarotation of these new sugars was not spontaneous, but depended on a process of "hydrocatalysis," *i.e.*, on the catalytic action of water, or on the action of other impurities which were capable of producing similar effects. This view was justified at the time by the fact that mutarotation in anhydrous solvents is intolerably slow; but it has since been vindicated conclusively by the discovery that the mutarotation of the sugars can be arrested completely by working under clean conditions in an anhydrous solvent and with a suitable sugar.§

The primary purpose of the present investigation was to discover a technique for producing at will the arrests of mutarotation which had previously been observed only sporadically. The ultimate object was to make use of the clean solutions prepared by this new technique in order to carry out a series of tests on the influence of traces of impurities on the velocity of mutarotation, since these tests could only be carried out satisfactorily when the solutions to which the impurities were added were so clean that no change of rotatory power was taking place in them. The former purpose has been to a large extent fulfilled, since it was found that "clean" solutions could be prepared with some approach to regularity in a silica flask which had been ignited in a furnace and allowed to cool in a desiccator. Under these conditions it was often possible to prove that the solution *in the flask* had remained unchanged overnight, or during a period of one or more days, although the samples usually underwent a more or less rapid mutarotation when transferred to a polarimeter tube for observation. Further experiments then showed that the addition of a drop of water to a flask containing a "clean" solution of tetra-acetylglucose in dry ethyl acetate was not sufficient to initiate the mutarotation, but that water containing a trace of acid or alkali was an extremely efficient catalyst; in these catalysed solutions, however, a maximum velocity of mutarotation was only developed after two or three hours, giving rise to inflected curves in place of the usual unimolecular type. These inflected curves, occurring with a sugar and solvent in which mutarotation is generally unimolecular, are an entirely new phenomenon, the occurrence of which is probably conditioned by the use of highly purified materials and clean apparatus. In the case of the alkaline catalyst, there is clear evidence that this peculiar action may be due to catalysis by the contaminated surface of the containing vessel. The progressive acceleration of the change might then be due to the development of some sort of chain-reaction, spreading from this surface into the interior of the liquid.

---

* Lowry, 'J. Chem. Soc.,' vol. 75, p. 220 (1899).

† 'Trans. Chem. Soc.,' vol. 85, p. 1049 (1904).

‡ Lowry, 'Trans. Chem. Soc.,' vol. 85, p. 1566 (1904).

§ See especially Lowry and Richards, 'J. Chem. Soc.,' vol. 127, p. 1385 (1925).

EXPERIMENTAL.

1. *Apparatus.*

(i) The *polarimeter* was the large instrument which has been used during the past 20 years for experiments on rotatory dispersion.*

(ii) The *polarimeter tube*, 2 dm. in length, was made of silica and provided with a silica stopper. The end-plates were usually of glass, although a pair of compensated quartz plates was sometimes used. Before use, the tube was washed out with absolute alcohol, then five times with distilled water, twice with chromic acid, and finally six times with water. It was dried in the same way as the distillation apparatus described below, but was not heated from the outside: before use it was kept in a desiccator over phosphoric oxide. This elaborate cleaning, however, left the tube in a very active condition, and much better results were obtained when a succession of clean solutions was examined in the tube without any intermediate process of cleansing. On the other hand, a *graduated silica flask*, which was heated to redness in a small furnace after being cleaned, and was then allowed to cool in a desiccator over phosphoric oxide, was usually found to be in a very satisfactory condition of cleanliness.

---

* 'Phil. Trans.,' A, vol. 212, p. 261 (1912).

(iii) The *source of light* was an enclosed mercury arc, and all the readings were made with the green line Hg 5461. In addition to its convenience, the use of an enclosed arc renders possible a degree of cleanliness that cannot be secured when a sodium flame is used as a source of light. Some special experiments showed that the light from the arc was without influence on the course of mutarotation. No difference in the readings was observed when the sugar was exposed to the full radiation of the arc (broadside on, through a silica water-jacket) for 10 minutes; moreover, the velocity of mutarotation was not altered by the use of light-filters, nor was it different before and after exposure to light.

2. *Materials.*

The experiments were all carried out with solutions of tetra-acetylglucose in dry ethyl acetate. The purified materials were prepared as follows :—

(i) *Purification of phosphoric oxide.*—Commercial phosphoric oxide was found to develop a smell of phosphine when allowed to stand in moist air. Since it was possible that phosphine might be an active catalyst, the lower oxides of phosphorus were removed from the pentoxide by ozonisation and sublimation. This purification is rendered difficult owing to the formation at about 400° (or at a lower temperature when the oxide is impure) of a " glass " which has no definite melting-point and only a small vapour pressure.* In order to purify the oxide, whilst avoiding the formation of the glass, oxygen from a cylinder was passed through a glass ozoniser worked by a 1-inch spark coil, and then over crude phosphoric oxide contained in a long tube. The tube, which was provided with ground inlet and outlet joints, lubricated with syrupy phosphoric acid, was heated gently at some distance above the burners of a combustion furnace, and the oxide was stirred occasionally by rotating the tube about these joints. Under these conditions the lower oxides were converted into the pentoxide, whilst the vapour of the pentoxide condensed in a long glass tube ground to one end of the heating tube. Great care must be taken that the oxide is heated to such a point that sublimation proceeds at a convenient speed, but without any formation of glass, since not only is the glassy form useless on account of its low vapour pressure, and irrecoverable, but it also attacks both glass and silica, with the result that the tube is cracked. After a sufficient quantity of the oxide had been ozonised and sublimed, the light feathery sublimate was swept from the condensation tube into the storage bottle. On testing the product for the presence of the lower oxide, by heating with mercuric chloride solution, only a barely perceptible precipitate was produced after boiling for five minutes.

---

* Smits and Rutgers, 'J. Chem. Soc.,' vol. 125, p. 2573 (1924).

(ii) *Preparation of dry ethyl acetate.*—Ethyl acetate from stock, or specially prepared in the laboratory, was washed with dilute sodium carbonate, allowed to stand over solid potassium carbonate for a day, and then shaken mechanically with successive quantities of calcium chloride, and finally with purified phosphoric oxide. The ester was then fractionated twice over phosphoric oxide in an all-glass still, and finally distilled from the carefully dried still without the oxide. This method appears to produce a very pure and dry ethyl acetate, since, although a certain amount of charring occurs on distillation, the liquid distils within one-tenth of a degree. The purified solvent was kept in a silica flask which had been heated to redness and allowed to cool in a vacuum desiccator containing phosphoric oxide.

The distillation-apparatus consisted of a flask, fractionating-column and condenser, united by ground glass joints. The fractionating column was filled with half-centimetre lengths of glass tube, and the thermometer was totally enclosed in the system. Before use, the apparatus was filled with distilled water and left for a few hours; it was then emptied, drained, filled with chromic acid, and left for four hours. After washing away the chromic acid, the apparatus was again filled with water and left for a few hours to soak. No alcohol or ether was used for drying. This was carried out by passing a current of hot, dry, dust-free air, until the water had been removed: the apparatus was then heated as far as was considered to be safe, whilst air was still passing through it. It was finally heated whilst under the vacuum of a water-pump. Although this procedure is not so drastic as baking out for several hours under the highest vacuum of a diffusion-pump, and at a temperature just below the annealing temperature of the glass, yet it is certain that not much organic matter remained adsorbed on the walls and that the apparatus was substantially clean.

(iii) *Desiccation of the sugar.*—A supply of tetra-acetylglucose was very generously placed at our disposal by Messrs. Boots Pure Drug Company.* Before each experiment the sugar was dried for three days in a special desiccator, the interior of which was maintained (during the day) at a temperature of 35°. The desiccator contained purified phosphoric oxide. When the sugar was required for use, air was admitted through a sulphuric acid bubbler, a glass wool filter, and a tube of phosphoric oxide. The dry sugar was then taken out of the desiccator, and a portion transferred to the silica flask as quickly as possible, and the solvent added. The sugar was not weighed except in an experiment carried out to determine its specific rotation. The exposure of the components to the moist atmosphere was thus very short. Since the object of the work was to note the conditions for the production of arrests, the concentration of the solutions could be deduced with sufficient accuracy from the rotations, the absolute values of which were known from an experiment carried out expressly in order to determine them. The assumption of a proportionality between concentration and rotation was justified by the fact that the concentrations were all of the same order of magnitude, viz., 5 to 10 per cent.

---

* Compare 'J. Chem. Soc.,' p. 722 (1926).

(iv) *Preparation of ketene.*—In some of the experiments a supply of ketene, $CH_2 : CO$, was required. This was prepared from acetone by the action of a hot wire in an apparatus (fig. 1) similar to that used by Wilsmore,* but

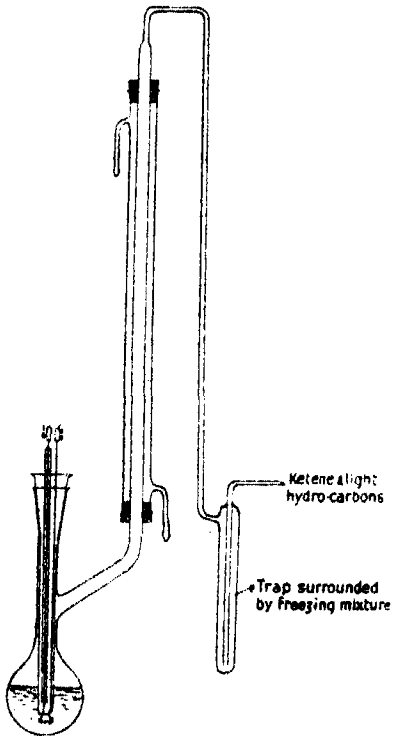

Ketene & light hydro-carbons
Trap surrounded by freezing mixture

FIG. 1.—Apparatus for preparation of ketene by Wilsmore's method.

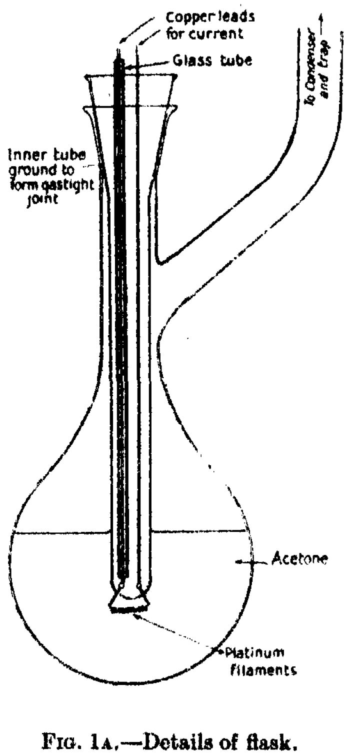


FIG. 1A.—Details of flask.

---

* 'Nature,' vol. 75, p. 510 (1907).

constructed with a ground-glass stopper instead of a rubber stopper protected by a layer of mercury.

To prepare the ketene, acetone was placed in the flask and direct current from the 110-volts supply was turned on gradually and regulated to a maximum of 4–6 ampères. If the current is brought to its final value too suddenly, the platinum filament fuses. A trap was provided for freezing out impurities with carbon dioxide and ether, so that only a mixture of hydrocarbons and ketene passed forward. In a typical experiment, the mixture contained 7·1 per cent. of ketene, whereas Wilsmore obtained a concentration of 10 per cent. by employing acetic anhydride.

3. *Experimental Results.*

Since each experiment was carried out for a different purpose, the results cannot readily be summarised, and are therefore described consecutively.

Experiment 1.—*Arrest of mutarotation in a silica flask* (fig. 2).—This experiment was carried out in order to show that the velocity of mutarotation of a solution of tetra-acetylglucose in ethyl acetate, instead of having a finite value which remained constant on the addition of small quantities of water,* could be reduced to zero if the sugar and solvent were properly purified, and examined under clean conditions. The material used for this purpose was a small specimen (less than a gramme), which had been given to us by Prof. Haworth. After the sugar had been dried for some hours in a vacuum desiccator over phosphoric oxide, it was dissolved in dry ethyl acetate in an ignited silica flask. A portion was transferred to a silica polarimeter tube at 20°, and gave the rotations shown in Table I (*a*). After 24 hours, when the experiment appeared to have failed, the tube was emptied, rinsed with a little of the solution from the silica flask, and again filled with that solution. The rotations then observed are shown in Table I (*b*).

It will be seen that there had been a complete arrest of mutarotation *in the flask* during a period of 24 hours, since the observed rotation, instead of increasing by 2·41° (as in the tube), actually showed a minute apparent decrease of 0·02°. Moreover, whilst the solution in the tube gave an average velocity-coefficient of 0·0015 (about one-tenth as great as for glucose in water), a second sample transferred to the same tube showed a velocity-coefficient which was at least 10 times smaller than before, showing that the cleanliness of the tube had been improved enormously by mere contact with a clean solution (see fig. 2). In each case, however, there was a tendency for the mutarotation to

---

* Baker, Ingold and Thorpe, 'J. Chem. Soc.,' vol. 125, p. 268 (1924).

Table I.—Mutarotation of Tetra-acetylglucose in Dry Ethyl Acetate (0·5 g./100 c.c. at 20°).

	Time.	Rotation.*	Velocity-coefficient.	
	Min.	°		
(*a*)	10	0·090	—	Half-change period = 0·4 day.
	20	0·130	0·0012	
	35	0·180	0·0012	
	50	0·240	0·0013	
	85	0·400	0·0015	
	175	0·806	0·0018	
	245	1·040	0·0019	
	275	1·142	0·0018	Half-change period = 0·25 day.
	335	1·300	0·0018	
	635	1·940	0·0018	
	1435	2·500	0·0015	
		Mean	= 0·0015	
(*b*)	1455	0·073	—	
	1475	0·080	0·00013	
	1595	0·146	0·00021	
	3110	0·487	0·00010	
	4870	0·659	0·00007	Half-change period = 7 days.
	5185	0·686	0·00007	
	7755	0·824	0·00004	
		Mean	= 0·00010	

---

* The mean of a group of readings to 0·01° on an instrument graduated to read to 0·001° is given to 0·001°. The last figure was used for calculation, but has very little absolute value.

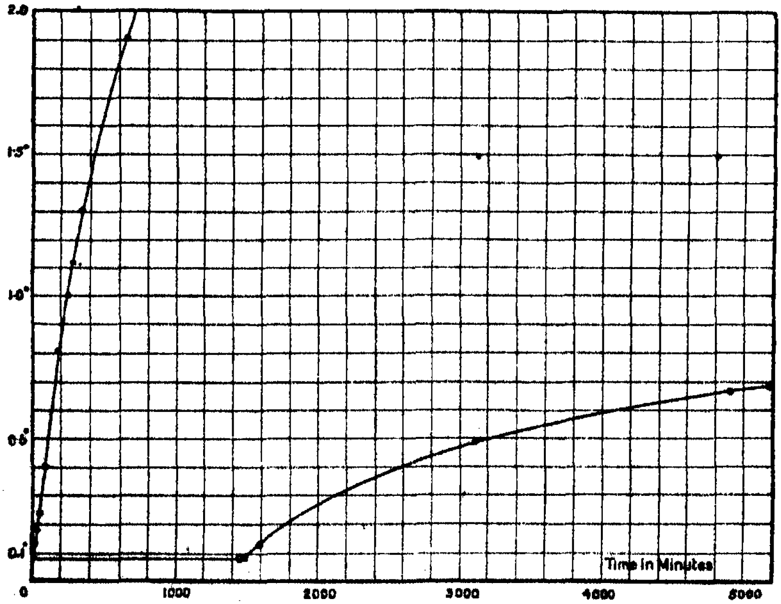


FIG. 2.—Mutarotation of tetra-acetylglucose in ethyl acetate (showing arrest of mutarotation in a silica flask).

accelerate and then retard, as if the solution were acquiring from the polarimeter-tube traces of a catalyst which gradually became ineffective.

Experiment 2.—*Progressive purification of a polarimeter tube.*—When a larger stock of sugar was available, a second experiment was made with a solution containing 7·5 grammes of the sugar in 100 c.c. of a solution in dry ethyl acetate, in order to diminish the relative magnitude of the experimental errors. The observed rotation of the solution *in the polarimeter tube* increased only by 0·08° (out of a total change of 13·7°) during the first six hours, and by a further 0·05° during the succeeding night. The velocity-constant for the solution *in the tube* was therefore 0·000024 (corresponding with a half-change period of 20 days), *i.e.*, only one-fourth of the velocity recorded in Experiment 1 (*b*), or one-sixtieth of that recorded in 1 (*a*).

Table II.—Mutarotation of Tetra-acetylglucose in Dry Ethyl Acetate (7·5 g./100 c.c. at 20°).

	Time.	Rotation.	Velocity-coefficient.	
	Min.	°		
(*a*)	8	0·859	—	
	14	0·875	—	
	21	0·859	—	
	31	0·868	—	
	40	0·863	—	
	52	0·862	—	
	72	0·875	0·000031	Mean = 0·000024. Half-change period = 20 days.
	139	0·882	0·000011	
	197	0·919	0·000019	
	222	0·925	0·000027	
	281	0·938	0·000024	
	376	0·938	0·000029	
	1305	0·990	—	
The solution in the flask was then examined.				
	Min.	°		
(*b*)	1341	2·617	—	
	Final	14·60	—	

Experiment 3.—*Influence of minute traces of water.*—This experiment showed strong retardation of mutarotation both in the polarimeter-tube and in the flask. The half-change period was about a week, but that of the second sample had increased at the end of 100 hours to about three weeks; a drop of ethyl acetate containing 1 per cent. of water was then added to it, giving a concentration in the tube of 0·002 per cent. H_2O. The addition of this minute trace of water was intended to make "Bakerian dryness" impossible, but without introducing enough water to act as a catalyst *per se*. No immediate change in the rate of mutarotation was observed, but the velocity-coefficient, which had diminished gradually from 0·000092 to 0·000017, increased again slowly from 0·000015 to 0·000027.

Table III.—Influence of Traces of Water on Mutarotation of Tetra-acetylglucose in Ethyl Acetate (6·82 g./100 c.c. at 20°).

Time.	Rotation.	Velocity-coefficient.
Min.	°	
(*a*) 4	0·830	—
8	0·837	—
20	0·841	—
34	0·847	—
93	0·942	0·000097
144	0·973	0·000080
221	0·031	0·000069 } Half-change
1285	0·122	0·000070 } period = 7 days.
The tube was re-filled with solution from the flask.		
Min.	°	
(*b*) 1302	2·33	—
1319	2·33	0·000092
1555	2·38	0·000082
1661	2·39	0·000075
1796	2·41	0·000071
2768	2·50	0·000052
4076	2·54	0·000034
5996	2·80	0·000017
One drop of a 1 per cent. solution of water in ethyl acetate was added to the solution in the tube.		
Min.	°	
(*c*) 7005	2·795	—
7011	2·795	—
7036	2·790	0·000015
8371	3·020	—
8716	3·020	0·000020
9811	3·403	—
11250	4·030	0·000027

Experiment 4.—This experiment was a failure, since mutarotation was almost complete, both in the tube and in the flask, on the day following that on which the solution was prepared.

Experiment 5.—*Influence of water on velocity of mutarotation* (fig. 3).—A very slow mutarotation, with a half-change period of 75 days, was recorded in the first sample poured into the polarimeter tube, the velocity-constant, 0·000006, being four times less than in Experiment 2 and 250 times less than in Experiment 1 (*a*). The solution *in the flask*, however, showed the same perfect arrest as in Experiment 1 (*b*), since there was again a minute apparent *decrease* of 0·02° in the rotatory power. This arrest persisted after the sample had been transferred to the tube over a period of an hour, at the end of which a drop of water was added to the solution in the tube, giving a concentration of 0·3 per cent. of water. After an initial fluctuation, a slow mutarotation was produced, but this gradually fell off until the half-change period had risen again to the high value of about 70 days, which was recorded for the dry sample first examined in the tube.

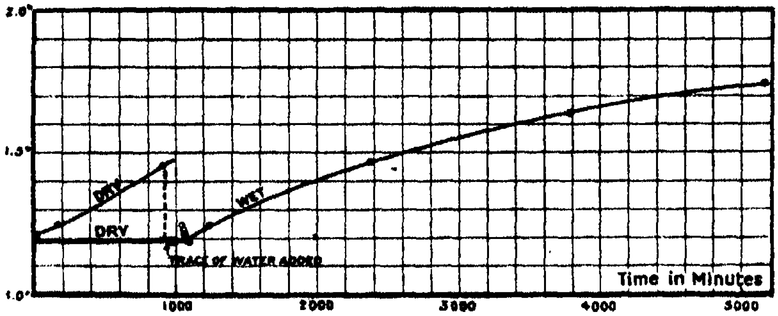


Fig. 3.—Influence of water.

Table IV.—Mutarotation of Tetra-acetylglucose in Dry and Wet Ethyl Acetate (9·75 g./100 c.c. at 20°).

	Time.	Rotation.	Velocity-coefficients.
	Min.	°	
(a)	11	1·210	—
	16	1·217	—
	156	1·247	0·0000065 Half-change period = 75 days.
	896	1·453	0·0000064
The tube was re-filled from the flask.			
	Min.	°	
(b)	935	1·192	An arrest has been produced in the flask, and persisted in the tube after the transfer of the solution.
	937	1·147	
	947	1·178	
	958	1·196	
	974	1·197	
	999		
One drop of water was added to the solution *in the tube*.			
	Min.	°	
(c)	1030	—	—
	1032	1·263	—
	1036	1·246	—
	1044	1·226	—
	1063	1·218	—
	1082	1·191	—
	1219	1·248	0·000022
	2359	1·469	0·000012
	3774	1·664	0·0000092
	5169	1·744	0·0000070

Experiment 6.—Velocity-coefficient in the tube 0·00012, in the flask 0·000086, half-change periods 4 and 5½ days.

Experiment 7.—*Use of ketene; impotence of water as a catalyst* (fig. 4).—This experiment showed marked points of similarity to the preliminary Experiment 1; but the ethyl acetate, instead of being purified by distillation from phosphoric oxide, was treated with a little ketene, in order to eliminate traces of water and of basic impurities, just as carbonyl chloride was used to arrest the mutarotation of nitrocamphor in benzene and ether. Since ketene is a very powerful acetylating agent, any surplus would spend itself quickly on the sugar, and would thus be eliminated before the measurements of the optical rotatory power of the solution had made any substantial progress. This conclusion was borne out by the experimental results, which showed that the "ketenised" solution in the flask was pure enough to give a change of rotatory power of only 0·11° in 828 minutes, but that the anti-catalyst which had been added was no longer available to combat the impurities in the polarimeter-tube, where the mutarotation was only 10 times less rapid than in an aqueous solution of glucose. The retardation of mutarotation in the flask persisted in an even more pronounced form after a drop of water had been added to it, since a further change of only 0·08° was recorded in the next 2304 — 828 = 1476 minutes, the velocity-coefficient being reduced to one-half after the addition of water. A progressive retardation of mutarotation was also observed in the polarimeter-tube, as successive samples were transferred to it from the flask, *including the sample to which a drop of water had been added*, which changed only about half as fast as the preceding sample. These experiments confirm the conclusion

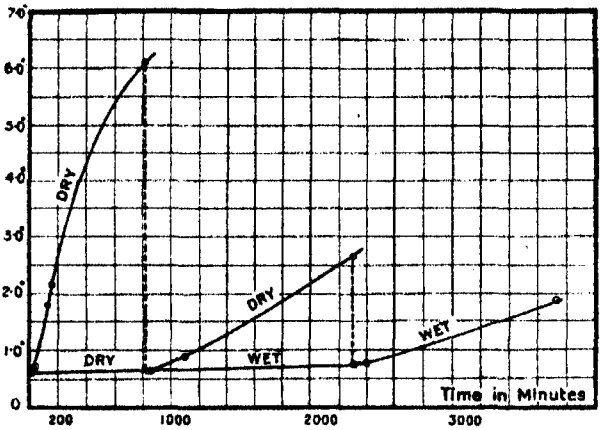


FIG. 4.—Influence of water.

that Bakerian purity rather than Bakerian dryness is needed in order to produce an arrest of mutarotation.

Table V.—Mutarotation of Tetra-acetylglucose in Dry and Wet Ethyl Acetate (428 g./100 c.c. at 20°).

Time.	Rotation.	Velocity-coefficient.
Min.	°	
(a) 3	0·531	—
9	0·517	—
20	0·582	0·00044
28	0·632	0·00053
33	0·694	0·00067
120	1·780	0·0014
162	2·160	0·0014
812	6·087	0·0014
(b) *At this point the tube was re-filled from the flask, and one drop of water was added to the remainder of the solution in the flask.*		
Min.	°	
828	0·638	0·00002 = velocity in flask without water.
837	0·634	—
1099	0·891	—
2209	2·638	—
(c) The tube was re-filled with the solution in the flask to which water had been added.		
Min.	°	
2304	0·720	—
2407	0·777	—
3742	1·854	0·00001 = velocity in flask with water.
5124	3·129	—
8402	5·185	—

Experiment 8.—*Dilute acid as a catalyst* (fig. 5).—A solution of tetra-acetylglucose in ethyl acetate to which a trace of ketene had been added

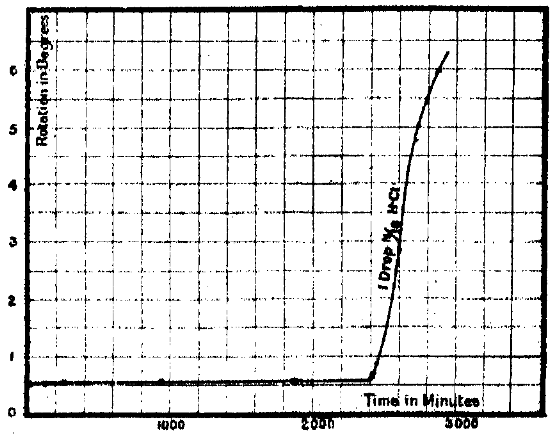


Fig. 5.—Influence of dilute acid.

gave an almost complete arrest of mutarotation *in the tube*, the rotation increasing by only 0·11° in 2388 minutes, giving a half-change period of about 112 days. Addition of a drop of N/10 HCl to the solution in the tube, however, caused a rapid mutarotation, which accelerated during a period of about 400 minutes, when the velocity-coefficient rose to 0·002. Thus, whereas in Experiment 7 a solution to which a drop of water had been added changed only half as fast as the dry solution, in the present experiment a solution containing about 2 per cent. of water, together with an acid-concentration of about 20 parts of HCl per million, produced a velocity of mutarotation that was only seven times less than that of glucose in pure water. In the meanwhile, the mutarotation of the solution in the flask was even more effectively arrested than during the preceding interval, since at 4525 minutes it had only changed by 0·19°, corresponding to a half-change period of 220 days.

Table VI.—Activation by Acid of Mutarotation of Tetra-acetylglucose in Ethyl Acetate (4·25 g./100 c.c. at 20°).

Time.	Rotation.	Velocity-coefficient.
Min.	°	
(*a*) 7	0·474	—
12	0·510	—
24	0·510	—
28	0·510	—
38	0·498	—
119	0·517	0·0000093 { Half-change period = 54 days.
927	0·563	0·0000084
2388	0·582	0·0000043 { Half-change period = 112 days.
One drop of N/10 HCl was added to the solution in the tube.		
Min.		
(*b*) 2393	°	
2398	0·592	—
2409	0·608	—
2416	0·668	—
2439	0·846	—
2596	2·800	—
2706	4·757	0·0017
2731	5·005	0·0020
2794	5·420	0·0024
2872	5·942	0·0021
4403	7·402	—
The tube was now filled with a fresh sample from the flask.		
Min.	°	
(*c*) 4525	0·662	0·0000027 { Half-change period = 220 days.
4557	0·666	—
4579	0·689	—

Experiment 9.—*Influence of dilute alkali* (figs. 6 and 7).—This experiment resembled Experiment 8, but the water added as a catalyst was activated with

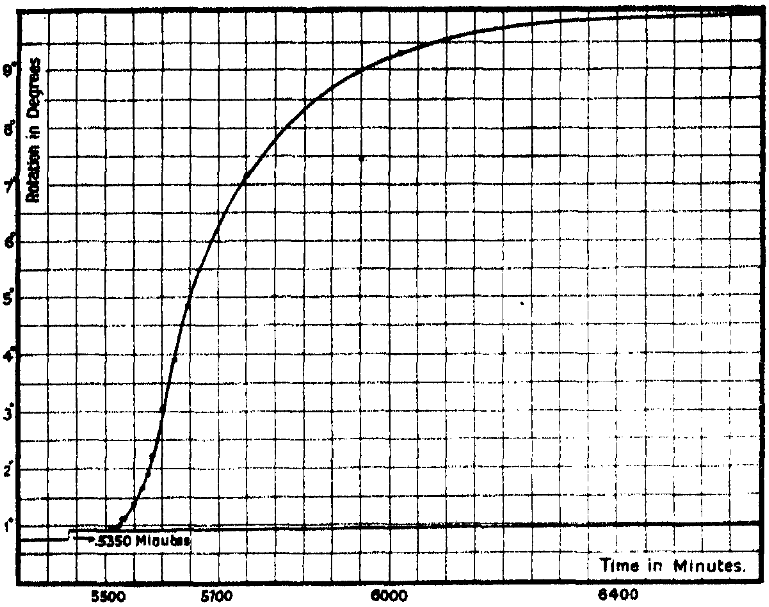


Fig. 6.—Influence of dilute alkali.

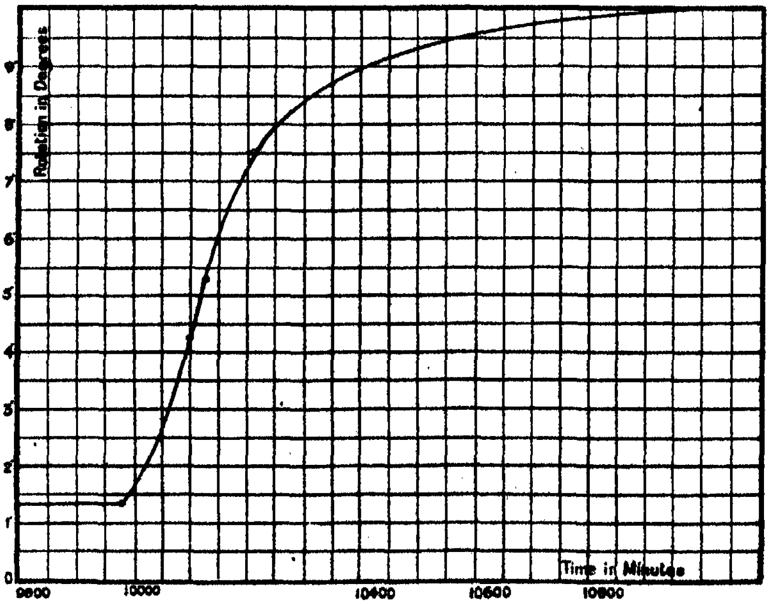


Fig. 7.—Influence of dilute alkali (continued).

alkali instead of with acid. At the end of 90 hours the clean solution in the *tube* showed a half-change period of about 96 days: one drop of N.NaOH was then added to it. A mutarotation was produced which accelerated gradually from 0·0015 to 0·0043, *i.e.*, from one-tenth to one-quarter of the velocity of mutarotation of glucose in water. At the end of about 160 hours, when the mutarotation of the first sample was completed, the tube was emptied and rinsed with the solvent: it was then filled with a second sample of the clean solution in the flask, which by this time had changed by 0·72° in 10,000 minutes (half-change period 70 days). Mutarotation started immediately, and pursued almost exactly the same course as after the original addition of alkali, although only a film can have been left on the inner surface of the tube.

Table VII.—Activation by Alkali.

Time.	Rotation.	Velocity-coefficient.	
Min.	°		
(*a*) 4	0·633	—	
7	0·647	—	
18	0·662	—	
40	0·683	—	
43	0·683	—	
86	0·666	—	
175	0·653	—	
290	0·663	—	
425	0·680	—	
1170	0·698	—	
1339	0·702	—	
1617	0·717	—	
2621	0·785	0·0000052	Half-change period = 96 days.
3077	0·818	0·0000055	
5512	0·938	0·0000053	

At 5518 min. one drop of N.NaOH was added to the solution in the tube.

Time.	Rotation.	Velocity-coefficient.
Min.	°	
(*b*) 5520	0·978	—
5524	1·122	—
5532	1·144	0·0015
5547	1·344	0·0015
5565	1·662	0·0017
5572	1·897	0·0019
5580	2·205	0·0025
5588	2·497	0·0025
5601	3·037	0·0029
5646	4·848	0·0040
5751	7·153	0·0043
5914	9·031	0·0043
6921	10·400	—
9946	10·530	—

Table VII—(continued.)

Time.	Rotation.	Velocity-coefficient.
At 9972 min. the tube was emptied, rinsed and filled again from the flask.		
Min.	°	
(*c*) 9978	1·349	0·0000069
10096	4·241	—
10126	5·344	—
10221	7·529	—
11411	10·323	—
12120	10·470	—
15120	10·618	—

4. *Summary and Conclusions.*

(*a*) By making use of an ignited silica flask, solutions of tetra-acetylglucose in dry ethyl acetate have been prepared which were sufficiently pure to show no change of rotatory power during a period of several hours.

(*b*) The addition of a drop of water to such a solution does not suffice to initiate mutarotation, but a rapid change is induced when a drop of dilute acid or alkali is added; the mutarotation curves are, however, inflected instead of uni-molecular, as if the action had been resolved into two consecutive stages.

(*c*) Since the action of alkali was the same when the polarimeter tube was rinsed with the solvent and re-filled with a second sample of the clean dry solution, it appears possible that these inflected mutarotation curves are characteristic of surface action, whereas the normal uni-molecular curves are developed by homogeneous catalysis.

Note on Total Reflexion of Electric Waves at the Interface between Two Media.

By H. M. MACDONALD, F.R.S.

(Received May 21, 1928.)

In a previous communication* it was assumed that, when total reflexion takes place at the interface between two media, the electric force in the disturbance in the second medium is in the plane of the wave-front; it may be shown that it is impossible in this case to satisfy the conditions that the electric and magnetic forces are both in the wave-front in the second medium. The object of the present communication is to investigate the disturbance in the second medium, and to obtain the changes of phase in the reflected waves in the first medium.

Taking the plane $z = 0$ as the interface between the two media, let the components of the electric force in the incident waves, $z > 0$, be given by

$$(A, B, C) \cos \kappa (lx + my + nz + Vt),$$

where

$$Al + Bm + Cn = 0,$$

the components of the electric force, when $z > 0$, are

$$\begin{aligned}(E_x, E_y, E_z) = (A, B, C) \cos \kappa (lx + my + nz + Vt) \\ + (A_1, B_1, C_1) \cos \kappa (lx + my - nz + Vt) \\ + (A_2, B_2, C_2) \sin \kappa (lx + my - nz + Vt),\end{aligned}$$

where, since

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0,$$

$$A_1 l + B_1 m - C_1 n = 0, \quad A_2 l + B_2 m - C_2 n = 0,$$

and, when $z < 0$,

$$\begin{aligned}(E_x', E_y', E_z') = (A', B', C') e^{\nu z} \cos \kappa (lx + my + Vt) \\ + (A'', B'', C'') e^{\nu z} \sin \kappa (lx + my + Vt),\end{aligned}$$

where, since

$$\frac{\partial E_x'}{\partial x} + \frac{\partial E_y'}{\partial y} + \frac{\partial E_z'}{\partial z} = 0,$$

$$\kappa (A'l + B'm) - \nu C'' = 0, \quad \kappa (A''l + B''m) + \nu C' = 0,$$

and

$$\nu^2 = \kappa^2 (\sin^2 \theta - \sin^2 \theta_0), \qquad \sin \theta_0 = (K'\mu'/K\mu)^{\frac{1}{2}},$$

---

* 'Roy. Soc. Proc.,' A, vol. 108, p. 386.

θ being the angle of incidence of the waves in the first medium, $\theta > \theta_0$, and K, μ; K′, μ' the constants of the first and second media respectively. The conditions that the components of the electric and magnetic forces parallel to the interface are continuous are

$$A + A_1 = A', \quad B + B_1 = B', \quad A_2 = A'', \quad B_2 = B'';$$

$$\sigma(nB - mC - nB_1 - mC_1) = -\nu B''/\kappa\mu'V - mC'/\mu'V,$$

$$-\sigma(nB_2 + mC_2) = \nu B'/\kappa\mu'V - mC''/\mu'V,$$

$$\sigma(lC - nA + lC_1 + nA_1) = lC'/\mu'V + \nu A''/\kappa\mu'V,$$

$$\sigma(lC_2 + nA_2) = lC''/\mu'V - \nu A'/\kappa\mu'V,$$

where $\sigma = (K/\mu)^{\frac{1}{2}}$. From the last four equations it follows that

$$\sigma n[(lB - mA) - (lB_1 - mA_1)] = -(\nu/\kappa\mu'V)(lB'' - mA''),$$

$$-\sigma n(lB_2 - mA_2) = (\nu/\kappa\mu'V)(lB' - mA'),$$

$$\sigma[n(Al + Bn) - (l^2 + m^2)C - n(A_1l + B_1m) - (l^2 + m^2)C_1]$$
$$= -(\nu/\kappa\mu'V)(A''l + B''m) - (l^2 + m^2)C'/\mu'V,$$

$$\sigma[n(A_2l + B_2m) + (l^2 + m')C_2]$$
$$= (l^2 + m^2)C''/\mu'V - (\nu/\kappa\mu'V)(A'l + B'm),$$

and eliminating C, C_1, C_2, C′, C″, the latter two equations become

$$\sigma[(Al + Bm) - (A_1l + B_1m)]/n = -(A''l + B''m)(\nu^2 - \kappa^2 \sin^2\theta)/\kappa\nu\mu'V,$$

$$\sigma(A_2l + B_2m)/n = -(A'l + B'm)(\nu^2 - \kappa^2\sin^2\theta)/\kappa\nu\mu'V,$$

whence, writing

$$\nu/\kappa\mu'V = \rho, \quad (\nu^2 - \kappa^2\sin^2\theta)/\kappa\nu\mu'V = -\kappa^2\sin^2\theta_0/\kappa\nu\mu'V = -\frac{\kappa K'V}{\nu} = -\frac{1}{\rho'},$$

and replacing A′, B′, A″, B″, by their values in terms of A, B, A_1, B_1, A_2, B_2,

$$(Bl - Am) + (B_1l - A_1m) = -(n\sigma/\rho)(B_2l - A_2m),$$

$$(Bl - Am) - (B_1l - A_1m) = -(\rho/n\sigma)(B_2l - A_2m),$$

$$(Al + Bm) + (A_1l + B_1m) = (\rho'\sigma/n)(A_2l + B_2m),$$

$$(Al + Bm) - (A_1l + B_1m) = (n/\rho'\sigma)(A_2l + B_2m),$$

therefore

$$B_1l - A_1m = (Bl - Am)(n\sigma/\rho - \rho/n\sigma)/(n\sigma/\rho + \rho/n\sigma),$$

$$B_2l - A_2m = -2(Bl - Am)/(n\sigma/\rho + \rho/n\sigma),$$

$$A_1l + B_1m = (Al + Bm)(\rho'\sigma/n - n/\rho'\sigma)/(\rho'\sigma/n + n/\rho'\sigma),$$

$$A_2l + B_2m = 2(Al + Bm)/(\rho'\sigma/n + n/\rho'\sigma),$$

and, writing

$$n\sigma/\rho = \tan \tfrac{1}{2}\chi, \qquad n/\rho'\sigma = \tan \tfrac{1}{2}\chi',$$

$$\mathrm{B}_1 l - \mathrm{A}_1 m = -(\mathrm{B}l - \mathrm{A}m)\cos\chi, \quad \mathrm{B}_2 l - \mathrm{A}_2 m = -(\mathrm{B}l - \mathrm{A}m)\sin\chi,$$

$$\mathrm{A}_1 l + \mathrm{B}_1 m = (\mathrm{A}l + \mathrm{B}m)\cos\chi', \quad \mathrm{A}_2 l + \mathrm{B}_2 m = (\mathrm{A}l + \mathrm{B}m)\sin\chi',$$

also, the last two are equivalent to

$$\mathrm{C}_1 = -\mathrm{C}\cos\chi', \quad \mathrm{C}_2 = -\mathrm{C}\sin\chi'.$$

It follows that

$$\mathrm{A}_1^2 + \mathrm{B}_1^2 + \mathrm{C}_1^2 + \mathrm{A}_2^2 + \mathrm{B}_2^2 + \mathrm{C}_2^2 = \mathrm{A}^2 + \mathrm{B}^2 + \mathrm{C}^2,$$

that is, the amplitude of the electric force in the reflected waves is equal to the amplitude of the electric force in the incident waves. The change of phase in the electric force in the plane of incidence parallel to the interface is $\pi - \chi$, and the change of phase in the electric force perpendicular to this direction and parallel to the interface is $-\chi'$. It may also be deduced that the components of the electric force in the second medium are

$$-2\mathrm{A}e^{\nu z}\sin\tfrac{1}{2}\chi\sin\{\kappa(lx + my + \mathrm{V}t) - \tfrac{1}{2}\chi\}$$
$$\qquad - 2\mathrm{C}e^{\nu z}\,nl\cos\tfrac{1}{2}(\chi - \chi')\cos\{\kappa(lx + my + \mathrm{V}t) - \tfrac{1}{2}(\chi + \chi')\},$$

$$-2\mathrm{B}e^{\nu z}\sin\tfrac{1}{2}\chi\sin\{\kappa(lx + my + \mathrm{V}t) - \tfrac{1}{2}\chi\}$$
$$\qquad - 2\mathrm{C}e^{\nu z}\,nm\cos\tfrac{1}{2}(\chi - \chi')\cos\{\kappa(lx + my + \mathrm{V}t) - \tfrac{1}{2}(\chi + \chi')\},$$

$$-(\kappa n/\nu)\,\mathrm{C}e^{\nu z}\sin\{\kappa(lx + my + \mathrm{V}t) - \chi'\}.$$

A Critical-Absorption Photometer for the Study of the Compton Effect.

By Profs. C. V. Raman, F.R.S., and C. M. Sogani, M.Sc.

(Received March 5, 1928.)

[Plate 10.]

1. *Introduction.*

That a change of wave-length occurs in X-ray scattering was first indicated by absorption measurements with the ionisation chamber, which showed that the absorption coefficient of a light element like aluminium was slightly greater for the scattered than for the primary X-rays. Later* more conclusive and direct evidence was obtained when spectrometric analysis of the scattered X-rays was made first by the ionisation and afterwards by the photographic method. This analysis disclosed the existence of an unshifted as well as the shifted line, and showed also that the latter becomes relatively more prominent with diminishing wave-length and lower atomic number of the scattering element. After the main features of the Compton effect were established by means of spectrometric measurements, however, absorption measurements with the ionisation method have again been employed for a detailed study of the phenomenon, for such measurements are much quicker than the spectrum experiments, where the final energy available is much smaller on account of the double scattering involved.†

As mentioned above, the absorption measurements were based on the slight increase in the absorption coefficient of a light element when the wave-length changes from the unmodified to the modified value. The much larger and sudden diminution in absorption of X-rays when the frequency is altered from the short to the long wave-length side of the critical K-absorption limit of the element used as a filter, furnishes us with an easy and convenient method of exhibiting the wave-length change in X-ray scattering. In the present paper will be described a photographic wedge photometer based on this principle, which enables the characteristics of the Compton effect to be readily observed. It may be pointed out that the same idea could no doubt be utilised also in connection with the ionisation measurements of the Compton effect.

---

* A. H. Compton, 'Phys. Rev.,' vol. 22, p. 408 (1923). P. A. Ross, 'Proc. Nat. Acad. Sci.,' vol. 10, p. 304 (1924). H. M. Sharp, 'Phys. Rev.,' vol. 26, p. 691 (1925). Y. H. Woo, 'Phys. Rev.,' vol. 27, p. 119 (1926), etc.

† De Foe, 'Phys. Rev.,' vol. 27, p. 675 (1926).

2. *Principle and Description of the Photometer.*

The application of the idea of critical absorption demands a proper choice of the wave-length and the absorber. In other words, the wave-length λ of the radiation used should be so related to the critical K-absorption limit λ_κ of the absorber, that whereas λ is slightly smaller than λ_κ, the wave-length λ_ϕ of the modified ray given by the well-known Compton equation

$$\lambda_\phi - \lambda = 0\cdot0243\ (1 - \cos\phi) \qquad (1)$$

becomes greater than λ_κ when the angle of scattering is larger than some convenient angle ϕ_0. Confining ourselves to the characteristic K-radiation alone, a reference to the tables shows several substances for the target for which suitable absorbers are available. In Tables I and II is given a list of elements of moderate atomic numbers suitable for use as targets and absorbers, and the minimum angle of scattering ϕ_0 at which the combination becomes effective for exhibiting the Compton effect.

Table I.

At. No. of Target.	Target.	Ab-sorber.	ϕ_0	
			K_{α_2}	K_{α_1}
			°	°
21	Sc	Ca	115	125
22	Ti	Sc	37	49
31	Ga	Cu	121	132
32	Ge	Zn	134	148
33	As	Ga	62	72
34	Se	Ge	49	60
35	Br	As	22	40
38	Sr	Br	131	141
40	Zr	Rb	93	104
41	Nb	Sr	81	90

Table II.

At. No. of Target.	Target.	Ab-sorber.	ϕ_0	
			K_{α_2}	K_{α_1}
			°	°
42	Mo	Y	55	60
44	Ru	Nb	34	49
45	Rh	Mo	23	42
46	Pd	Mo	103	114
48	Cd	Ru	80	91
49	In	Rh	74	85
50	Sn	Pd	59	71
51	Sb	Pd Ag	108 57	120 69
52	Te	Ag Cd	103 49	115 62

In the present experiments the radiation from a tin target of a metal Shearer tube operated at about 60 K.V. was employed, and a palladium foil was used as filter. The wave-lengths of the characteristic radiations K_{α_1} and K_{α_2} for tin are 0·490 and 0·494 A.U. respectively, whereas the critical K-absorption limit of palladium is 0·506 A.U.

Equation (1) then gives $\phi_0 = 71°$ for the K_{α_1} line and $\phi_0 = 59°$ for the K_{α_2} line. Consequently, as ϕ is increased from a small to a large angle, the transparency of the palladium foil will undergo an enormous increase on passing these values of ϕ_0. To be more definite, the transparency of a foil 0·1 mm. thick will be increased nearly four hundred times in this process. Further, provided $\phi > \phi_0$, the transmitting power will be seen obviously to depend upon the relative proportions of the modified and the unmodified scattering in any given case, increasing with the former.

For the measurement of the intensity of the X-rays transmitted by the foil in any given case, a comparative method is adopted, the thickness of a light element like aluminium or iron, which is equivalent as regards absorption to the palladium being determined. To be able to do this from a single exposure, a wedge of such element (in the present case steel) with a suitable gradient of thickness is exposed side by side with the foil, the photographic plate being placed behind the combination and almost in contact with it. The resulting photograph will show two patches corresponding to the foil and the wedge respectively (see figs. 1 and 2, Plate 10), and the point of the wedge which possesses the same transmission as the foil can be determined by inspection, or more accurately with a microphotometer. Obviously the point of equality would shift towards the thinner end of the wedge as the modified radiation having wave-length $\lambda_\phi > \lambda_\kappa$ increases in proportion. It may be added that, for 0·1 mm. of palladium, the equivalent thickness of steel would be 0·6 and 0·08 mm. respectively when the whole of the scattering radiation lies on the short or the long wave-length side of the critical limit. The two dots in the figure correspond to two holes drilled in the steel wedge to serve as reference marks. Both the foil and the wedge were fixed in a metal plate with a suitable rectangular aperture, which was so designed as to serve as a plate-holder also. Arrangements were made to keep the photographic plate always in the same position with respect to the steel wedge. In order to protect the apparatus from stray radiation, the experimental beam was allowed to enter it only through a long lead tube.

3. *Experimental Results.*

Observations were made with three radiators, namely, carbon, aluminium and iron at various scattering angles. The results are set forth in Tables III and IV, where *t* in the last column represents the thickness of the steel wedge at the point of equal transmission. Figs. 1 and 2 (Plate 10) correspond to the scattering from carbon at 60° and 90° respectively. Figs. 3 to 8 are the reproductions of the microphotometer curves corresponding to the scattering from carbon at the various angles. In these curves the right-hand part corresponds to palladium, and the zero of the galvanometer falls outside the

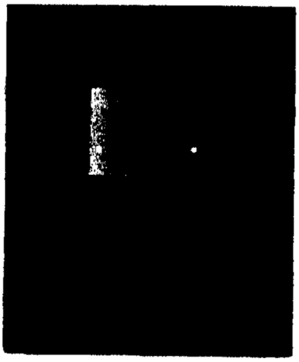


FIG. 1. $\phi = 60$

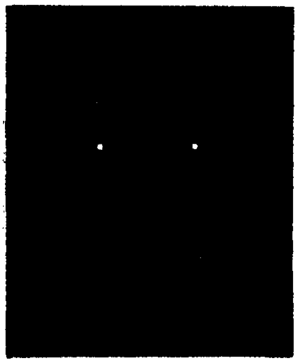


FIG. 2. $\phi = 90$

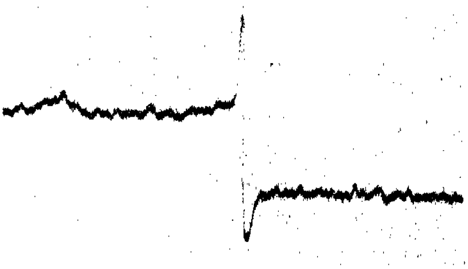


FIG. 3. $\phi = 45$

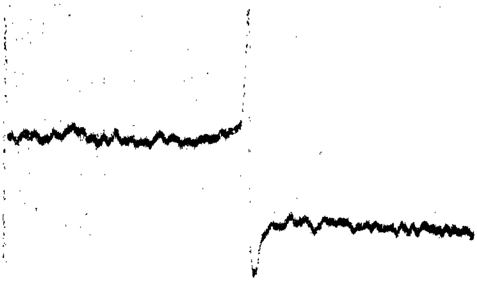


FIG. 4. $\phi = 60$

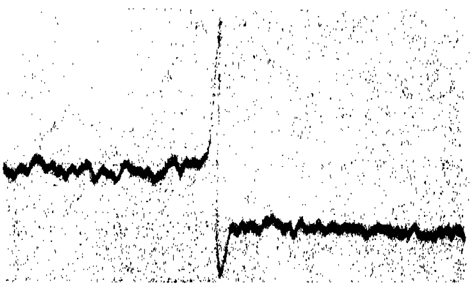


FIG. 5. $\phi = 75$

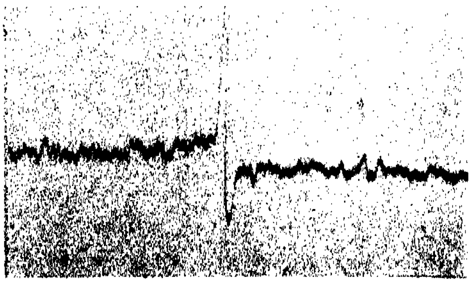


FIG. 6. $\phi = 90$

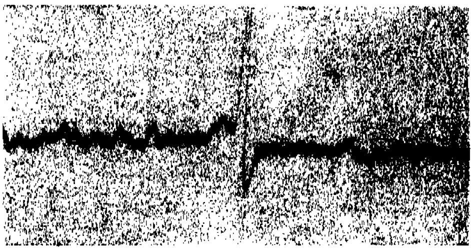


FIG. 7. $\phi = 105$

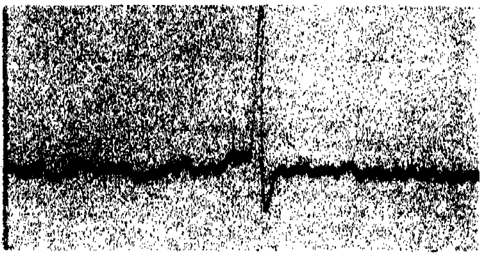


FIG. 8. $\phi = 120$

figures on the upper side. The microphotometer used was of the Moll type. For obtaining these intensity curves the negatives were run from top to bottom (figs. 1 and 2) across the same point of the wedge in each case.

Table III.

Radiator.	ϕ	t
	°	mm.
Carbon	45	0·32
"	60	0·31
"	75	0·27
"	90	0·22
"	105	0·22
"	120	0·22

Table IV.

Radiator.	ϕ	t
	°	mm.
Aluminium	45	0·36
"	60	0·36
"	90	0·33
"	120	0·33
Iron	60	0·35
"	90	0·34
"	120	0·34

In case of scattering from carbon a cylindrical rod of carbon about 2 cm. in diameter was irradiated so that the cross-section of the beam was a little greater than the diameter of the rod. To get the scattering from the more absorbing elements, namely, aluminium and iron, the primary beam was incident, nearly grazing the surface. The time of exposure for carbon was 10 minutes in each case, whereas exposures of 30 and 60 minutes were given in the case of aluminium and iron respectively. It may be remarked that the soft fluorescent radiation in the latter two cases would be cut off automatically to a large extent by the photometer itself.

From a survey of the observations given above the following characteristics of the Compton effect are brought out.

1. The change of wave-length of the modified line as a function of the scattering angle ϕ is given by equation (1). For, as calculated from that equation, the modified $K_{\alpha 2}$ and $K_{\alpha 1}$ remain on the short wave-length side of the critical absorption limit till angles of scattering 59° and 71° respectively are reached. Hence, the point of equality of transmission between the steel wedge and the palladium foil should remain practically fixed as the angle of scattering is increased till it reaches the value 59°. A considerable change at 71° is also to be expected. In agreement with these indications, the microphotometer records (figs. 3 to 6) show that there is practically no change till an angle 60° is reached, and that the most rapid change occurs between 60° and 90°.

2. The modified scattering is more prominent in the case of light radiators. Thus, while the change in the point of equality for carbon as we pass from

$\phi < \phi_0$ to $\phi > \phi_0$ is large enough to be clearly detected by the eye (compare figs. 1 and 2, Plate 10), the change is much less for aluminium and still less for iron.

3. The ratio of the modified to the unmodified radiation continues to increase with increased angle of scattering. This is clearly shown in the case of carbon by the sequence of the microphotometer records (figs. 5 to 8) for scattering angles 75°, 90°, 105° and 120°.

The reason why the arrangement does not show the expected sensitiveness is that there is a considerable amount of white radiation mixed with the characteristic in the case of heavy target like tin. In addition, there is the K_β also to be considered. The sensitiveness will therefore obviously be increased if suitable filters are employed for cutting off the K_β and most of the white radiation on either side of the K_α lines. This will, however, mean an increased exposure. A still better method would be to employ the radiation from a comparatively light element, in which case the white radiation will be relatively much less prominent. Even as it is, the apparatus is capable of readily exhibiting the main features of the Compton effect in an unambiguous manner.

4. *Summary.*

The great difference in transmission through a filter on the two sides of the K-absorption edge forms the basis of a very simple and convenient method of studying the characteristics of the Compton effect. For this purpose the material of the target for the X-ray tube and the material of the filter placed in the path of the scattered X-rays are so chosen that the unmodified scattered ray lies on the short-wave side of the limit, while the modified ray scattered through a sufficiently large angle is on the long wave-length side of the limit. The increased transparency for the scattered radiation at such angles is readily shown by a photographic plate exposed to the scattered radiation behind the filter, a steel wedge placed side by side with the filter forming the standard of comparison. The point of equality of transmission between wedge and filter shifts further and further towards the thin edge of the wedge with increasing proportion of modified to unmodified scattering. The known features of the Compton effect are unambiguously indicated.

The Emission of Secondary Electrons and the Excitation of Soft X-Rays.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society.

(Received May 7, 1928.)

Experiments I have recently made in collaboration with Dr. F. C. Chalklin* on the one hand and with Mr. F. S. Robertson† on the other, together with some observations not yet published made in the Wheatstone Laboratory by Mr. E. Rudberg,‡ taken in conjunction with Krefft's§ results for the secondary electron emission from baked tungsten, throw a good deal of light on the mechanism of the generation of secondary electrons at the surfaces of solids, particularly in the range where the energy of the primary electrons is sufficient to generate soft X-rays.

I shall first consider what are the chief essential facts from this point of view as to what happens when a beam of electrons falls on a conductor. In 1908|| I found that a considerable proportion of slowly moving electrons was reflected by a metal plate ; in the particular case of the electrons coming from a hot platinum strip under no applied electric force on to a brass plate, I estimated the proportion reflected at roughly 30 per cent. A similar result with electrons of 2, 4 and 8 volts, equivalent energy was obtained independently about the same time by von Baeyer.¶ Since that time a number of investigations of electron reflection at conductors have been published.** Speaking broadly, it appears that with increasing energy of the primary electrons the proportion reflected increases to a maximum, at a value which is in the neighbourhood of 11 volts for a number of metals, falls to a minimum at

---

* Richardson and Chalklin, 'Roy. Soc. Proc.,' A, vol. 110, p. 247 (1926); vol. 119, p. 60 (1928).

† Richardson and Robertson, 'Roy. Soc. Proc.,' A, vol. 115, p. 280 (1927).

‡ See, however, E. Rudberg, 'Nature,' vol. 119, p. 704 (1927).

§ H. E. Krefft, 'Ann. der Physik,' vol. 84, p. 639 (1927); 'Phys. Rev.,' vol. 31, p. 199 (1928).

|| O. W. Richardson, 'Phil. Mag.,' vol. 16, p. 898 (1908); vol. 18, p. 694; 'Phys. Rev.,' vol. 29, p. 557 (1909).

¶ O. von Baeyer, 'Verh. der Deutsch. Physik. Ges.,' 10 Jahrg., pp. 96, 953 (1908), and 'Phys. Zeits.,' 10 Jahrg., p. 168 (1909).

** For literature to 1922, see H. E. Farnsworth, 'Phys. Rev.,' vol. 20, p. 358 (1922). For later papers, see H. E. Krefft (§).

a value which is comparable to 30 volts, rises to a second maximum at a value which is of the order of 200 volts, and then falls off slowly and continuously with further increase in the energy. These results vary to some extent with the nature and treatment of the metal surfaces, but it is important to observe that there is generally some range of voltage in which the number of secondary exceeds the number of primary electrons. This is usually in the region in which the soft X-ray emission becomes important. Farnsworth* examined the electrons emitted from a nickel plate and found that with primary electrons having 9 volts energy or less, a large proportion of the secondary electrons had an amount of energy nearly equal to that of the primary electrons, and but a small proportion had a velocity of 1 volt or less. As the energy of the primary electrons was increased above 9 volts, the proportion of low velocity electrons steadily increased and the proportion of secondary electrons having energy close to that of the primary electrons steadily fell to a very small percentage at 110 volts, a result previously obtained by Davisson and Kunsman.†

These and other observations indicate that at low voltages of the order of 10 or less, the main effect is purely a reflection of the incident beam of electrons ; that at higher voltages a pure reflection of the type discovered and investigated by Davisson‡ and his associates persists ; but that this reflection gives rise to a continuously diminishing proportion of the secondary electrons as the energy of the primary electrons is increased.

The question that I am now chiefly concerned with is the mode of origin of this increasing proportion of secondary electrons which does not appear to result from reflection. Krefft's experiments show that close to each primary voltage at which there is a sudden change in the soft X-ray emission from tungsten there is also a similar sudden change in the secondary emission. This shows that the excitation of these secondary electrons and the excitation of soft X-rays is very closely associated. There are a number of subsidiary facts which in my judgment establish this close association very convincingly. In the first place, Krefft's secondary electron emission inflections occur at our uncorrected, not at our corrected voltages. This is shown by Table I, which contains all the inflections between 29 and 420 volts found by Krefft and all the tungsten inflections given by Richardson and Chalklin in their first paper, with the addition of two "impurity" breaks 137·6 (142·6) and

---

* *Loc. cit.*

† Davisson and Kunsman, 'Science,' vol. 64, p. 522 (1921); 'Phys. Rev.,' vol. 22, p. 242 (1923).

‡ Davisson and Germer, 'Nature,' vol. 119, p. 558 (1927), and 'Phys. Rev.,' vol. 30, p. 705 (1927).

Table I—(Volts).

Krefft, secondary electron emission	29·5	33·0	38·7	60·3	64·8	69·3	74·8	109·2	127·7	140
R. and C., uncorrected	31·9	35·4	40·0		65·9			103·3	122·8	137·6 145
R. and C., corrected	36·8	40·3	44·9		70·8			108·2	127·7	142·6 150

Krefft, secondary electron emission	160*	190*	237·5	249·0	257·5	330*	390*	415
R. and C., uncorrected	162	186	240·3		258·8	331	390	414
R. and C., uncorrected	167	191	245·2		263·7	336	395	419

162 (167) from their second paper. The only apparent exceptions are the soft X-ray inflections at 108·2, 127·7 and 191. Of these, only 108·2 is likely to be due to tungsten. This is a very weak break and there may be a considerable error in it. Furthermore, the recent measures of 191 have put it at about 197, and in the earlier work it was very variable. The remainder of the (uncorrected) soft X-ray breaks lie on an average 1·6 volts above the secondary electron emission breaks. If this shift were made, all the determinations, with the three exceptions noted, would be identical within the errors of measurement, and I am not disposed to claim that the absolute values of many of these discontinuities are sure to more than about a volt. It is, of course, essential that the coincidence should be at the uncorrected and not at the corrected soft X-ray values if the secondary electron emission and soft X-ray excitation are to be simultaneous effects. It is also important that the total number of discontinuities (20) is not large considering the range of voltage (nearly 400) covered. The average distance between the inflections is more than 20 times the probable error of most of them. As a matter of fact, a similar correspondence can be traced between Petry's[†] earlier secondary electron emission inflections for iron and the soft X-ray inflections, but with this element the number of inflections is so great that the correspondence might be a matter of chance and can only be looked upon as a confirmation of the inference which has been drawn from the tungsten results. It is satisfactory that Krefft has very few inflections which do not correspond with the soft X-ray inflections, no more, in fact, than might reasonably be expected to arise from impurities in the different samples of metals under investigation. Finally, the fact that Krefft found that the inflections marked * disappeared at high temperatures, and we had found that these and some related inflections could be produced

---

[†] R. L. Petry, 'Phys. Rev.,' vol. 26, p. 346 (1925).

by deposition and removed independently of the others must remove any remaining doubt there may be about this matter.

Although the evidence described above establishes the simultaneity of the production of characteristic soft X-rays and of secondary electrons, it still leaves possible a considerable amount of latitude as to the precise mechanism of the two emissions. This can only be settled by a discussion of the magnitudes of the quantities concerned. I shall start this by setting out the following list of ways in which the non-reflected secondary electron emission might possibly and not unreasonably be expected to originate, without doing violence to the facts which have been recapitulated :—

(*a*) The primary electrons interact with a large number of electrons in the conductor, and so, on the whole, they give the ejected electrons a very small velocity compared with their own.

(*b*) The first step is the absorption of a primary electron, energy eV, by the conductor with the conversion of its energy into a quantum of soft X-ray radiation of frequency given approximately by $\nu = eV/h$. The soft X-rays are converted into softer radiations by electronic interactions within the atoms as ordinary X-rays are. These softer radiations then eject electrons by photo-electric action.

(*c*) The primary electron of energy eV is first converted into a soft X-ray of frequency approximately $\nu = eV/h$ as in the first part of (*b*). This radiation ejects an electron by photo-electric action with energy comparable with eV. This electron causes the ejection of a number of electrons having smaller energies by the same process as in (*a*).

(*d*) The primary electrons by gradual decelerations without stoppage give rise to softer radiations than that of frequency about $\nu = eV/h$, and these eject slow electrons by photo-electric action.

Now Rudberg has tried the experiment of placing a thin window of good clear fluorite in the path of a typical beam of soft X-rays of the kind dealt with at the higher voltages in our experiments. He finds it cuts down the photo-electric emission at the testing plate to about 0·2 per cent. of its previous value. This shows that practically the whole of the soft X-rays dealt with in our experiments really are soft X-rays and not low voltage radiations of the ultra-violet type which could pass through fluorite. On the other hand, his analysis of the energy of the photo-electrons generated by these soft X-rays shows that nearly all of them have energy which is very small compared with the energy of a suitable soft X-ray quantum as judged, on the $h\nu = eV$ relation, from the voltage applied to the X-ray generating tube. The distribution is, in fact, quite similar to that found by Farnsworth for the secondary electron emission if reasonable allowance for the Davisson reflexions is made.

It seems then, that although energies equivalent to voltages below 10 are very common, both in the secondary electron emissions and in the photo-electrons from soft X-rays generated by primary electrons with energies of the order of a few hundred volts, radiations equivalent, on the $h\nu = eV$ relation, to voltages under 10 are quite rare in the soft X-ray beam. This, unless some modifying circumstances can be brought into account, would make untenable schemes (*b*) and (*d*), and, in fact, any schemes in which the slow secondary electrons, that is to say, any considerable proportion of them, result from the photo-electric action of radiations of a corresponding voltage on the $h\nu = eV$ relation.

We are thus left with only (*a*) and (*c*). These do not appear to be mutually exclusive ; in fact, if (*c*) can occur it would seem that (*a*) must be possible also ; (*a*), however, will not account for the correlation between the soft X-ray and the secondary electron discontinuities ; it does not even mention soft X-rays. It seems, therefore, that (*c*) or some other scheme which involves soft X-rays must be important.

In the case of secondary electron emission, the ratio of the number of secondary to the number of primary electrons is always of the order of unity, in the range of voltage with which we are concerned. In the case of soft X-rays, Richardson and Robertson concluded that the value of $\alpha N'/2\beta$ was between three and seven for all the 14 elements examined, which ranged in atomic number from 6 to 79. Here N' is the reduced effective atomic number which governs the collisions of the electrons in the generation of X-rays, according to an adaptation of Kramers' theory, α is the number of electrons generated at the photo-electric target per quantum of soft X-rays incident on it and βeV is the average energy of a soft X-ray quantum generated with a primary voltage V on the X-ray tube. On scheme (*c*) we expect 2β to be comparable with unity. Now the values found for $\alpha N'/2\beta$ are close to what we should anticipate for the values of N'. From this it would follow that α also is of the order unity.

Thus this argument leads to the conclusion that just as in the reflection of primary electrons there is approximately one secondary electron ejected for each primary electron incident so also in the photo-electric action of soft

X-rays there is approximately one electron ejected for each quantum of X-rays incident.

In each case, however, the average energy of the ejected electron is much less than that of the projectile which is the primary cause of its ejection.

Furthermore, the proportion of new soft X-rays generated in the neighbourhood of a given inflection, while larger, is of the same order of magnitude as the proportion of new secondary electrons. Thus, to take the well-marked tungsten inflection at about 265 volts, I find from fig. 8 of Richardson's and Chalklin's first paper that the increase of the ordinate in the 15 volts up to 280 is 0·2. In the same units the amount of other radiation is 1·75, giving a proportion of 0·114. In the case of the generation of secondary electrons, taking the corresponding inflection in fig. 9 of Krefft's paper the rise in 15 volts is about 1·6, the other secondary radiation in the same units being 109·6. This gives a proportion of 0·0146.

It is convenient at this stage to make a calculation of the expected intensity of a secondary beam which is generated when a primary beam falls on an absorbing substance bounded by a plane surface. The essential properties I assume for the beams are that they obey an exponential law of absorption with distance, and that the rate of conversion at any depth is proportional to primary intensity. Let the intensity of the primary beam before incidence be I_0, then, if it is incident at an angle θ to the normal, its intensity I at depth x is $I = I_0 e^{-\frac{K_1 x}{\cos\theta}}$, where K_1 is the absorption coefficient. The amount crossing area dS at depth x in unit time is

$$I \cos\theta\, dS = I_0 \cos\theta e^{-\frac{K_1 x}{\cos\theta}} dS. \tag{1}$$

These give rise to $\alpha I_0 \cos\theta e^{-\frac{K_1 x}{\cos\theta}} dS\, dx$ secondary rays, where α is the coefficient of conversion, which have an intensity

$$I' = \alpha I_0 \cos\theta e^{-\frac{K_1 x}{\cos\theta}} e^{-\frac{K_2 x}{\cos\phi}} dx \tag{2}$$

on emergence at an angle ϕ with the normal, K_2 being the co-efficient of absorption of the secondary rays. The incident beam being taken to have the properties of a plane wave, or a rectilinear projectile, and the secondary to be isotropic, we have for the total secondary intensity in the conical range of directions ϕ to $\phi + d\phi$

$$\tfrac{1}{2}\alpha I_0 \cos\theta \sin\phi\, d\phi \int_0^\infty e^{-\left\{\frac{K_1}{\cos\theta} + \frac{K_2}{\cos\phi}\right\} x} dx = \tfrac{1}{2}\alpha I_0 \cos\theta \sin\phi\, d\phi \Big/ \left(\frac{K_1}{\cos\theta} - \frac{K_2}{\cos\phi}\right). \tag{3}$$

If we want the total emission for all values of ϕ it is—

$$\tfrac{1}{2}\alpha I_0 \cos\theta \int_0^{\pi/2} \frac{\sin\phi\, d\phi}{\left\{\dfrac{K_1}{\cos\theta} + \dfrac{K_2}{\cos\phi}\right\}}$$

$$= \tfrac{1}{2}\alpha I_0 \frac{\cos^2\theta}{K_1}\left\{1 - \frac{K_2\cos\theta}{K_1}\log\left(1 + \frac{K_1}{K_2\cos\theta}\right)\right\}. \qquad (4)$$

This expression can be shown to be positive for all real values of $K_2 \cos\theta/K_1$, which is necessary. In considering these conversions, it is, perhaps, more convenient to regard both electrons and radiation as particles. A calculation similar to the foregoing shows that if N_0 is the number incident, α the number converted per unit length of track and N the total number which escape

$$N/N_0 = \frac{\alpha\cos^2\theta}{2K_1}\left\{1 - \frac{K_2\cos\theta}{K_1}\log\left(1 + \frac{K_1}{K_2\cos\theta}\right)\right\}. \qquad (5)$$

We shall now consider what is known about K_1 or K_2 for electrons of the energies here dealt with and for soft X-rays. Some data given by Lenard lead to a value of K_1 for copper for 300 volt electrons of about 6×10^5 cm.$^{-1}$. Some calculations by Partzsch and Hallwachs* give for electrons liberated in gold by ultra-violet light K_2 about 10^6 cm.$^{-1}$. These would be electrons moving with energy equivalent to about 6 volts. Lenard's results indicate that at these low voltages the diffusion loss from an electron stream does not vary very rapidly with the voltage. This can, therefore, be regarded as confirming the approximate accuracy of the figure got from Lenard's data. The value for substances other than copper can be got roughly by assuming K_1 to be proportional to the density. The density of copper is 8·93. For soft X-rays at 350 volts some direct measurements of absorption made in the Wheatstone laboratory by M. Bandopadhyaya give for aluminium $K_1 = 8{\cdot}75 \times 10^4$ cm.$^{-1}$ and for gold $K_1 = 3{\cdot}76 \times 10^{+5}$ cm.$^{-1}$. For carbon K_α (280 volt) radiation, using Holweck's† oxygen and nitrogen absorption measurements, and assuming that the absorption varies as $\lambda^{5/2}$, and as the square of the effective atomic number I find that the value of K_1 for carbon K_α radiation in copper is about 10^5. This may be regarded as a confirmation of Bandopadhyaya's measures.

These figures lead to the conclusion that the absorption coefficients K_1 are very similar both for soft X-rays and for the electrons which produce them. The values may not be very exact, but they preclude the possibility that these quantities have totally different magnitudes. It is natural to assume that for the *same radiation* K_1 and K_2 have the same value in equation (5).

---

* Partzsch and Hallwachs, 'Ann. d. Physik,' vol. 41, p. 269 (1913).

† Holweck, 'Ann. de Chim. et de Phys.,' vol. 17, p. 40 (1922).

Since we are now only concerned with the order of magnitude of the quantities involved we can neglect the obliquity factor cos θ in (5). Furthermore, unless K_2 and K_1 have totally different magnitudes the factor

$$1-\frac{K_2\cos\theta}{K_1}\log\left(1+\frac{K_1}{K_2\cos\theta}\right)$$

is of the order of magnitude of unity. It thus appears that in all the cases with which we are concerned N/N_0 is comparable with $\alpha/2K_1$. In the secondary electron emission N/N_0 is of the order unity, so that α_1 the number of times a primary ejects a secondary electron per cm. of its path is comparable with $2K_1$, and therefore of the order 10^6. In the excitation of soft X-rays N/N_0, the number of X-ray quanta, divided by the number of incident electrons is about 10^{-6}; so that α, the number of electrons converted into an X-ray quantum per unit length of electron track, is comparable with $2\times10^{-6}K_1$ or unity.

This leads to a difficulty in connecting Krefft's results for secondary electron emission with the excitation of soft X-rays. If a primary electron produces 10^6 times as many secondary electrons as X-ray quanta per cm. of its path, then we expect the secondary emission which arises by the intervention of soft X-rays to be only of the order of 10^{-6} of the total observed secondary emission. But Krefft's results require that while the proportion of the secondary emission which is attributable to a given inflection is a somewhat smaller proportion of the total secondary emission than the corresponding increment in soft X-ray emission is to the total soft X-ray emission, it is, nevertheless, of the same order of magnitude. On the other hand, the total secondary emission is comparable with the primary beam, whereas the total soft X-ray emission is only about 10^{-6} of the strength of the primary electron beam which excites it.

This difficulty can be avoided if we suppose that K_2 for soft X-rays generated by electron impact in a substance is very much higher than K_1 for the same soft X-rays as measured by transmission experiments. The difference must be very great, K_1 being of the order 10^6 cm.$^{-1}$ and K_2 at least 10^{10} cm.$^{-1}$.

This means that the reabsorption of the soft X-rays and conversion into electron streams close to the atom at which they originate is almost a certainty. The equation (5) with cos $\theta=1$ gives $N/N_0=\alpha/2K_1$, for K_2/K_1 small, N/N_0 of the order of $\alpha/2K_1$, when K_1 and K_2 are comparable and gives

$N/N_0 = \alpha/4K_2$ when K_2/K_1 is large. We thus have the following cases, using very rough magnitudes:—

Soft X-Ray Excitation.—$K_1 \approx 10^6$ $K_2 \approx 10^{12}$ $\alpha \approx 10^6$,

$$N/N_0 \approx \alpha/4K_2 \approx 10^{-6}.$$

Soft X-Ray Photo-Electric Action.—$K_1 \approx 10^6$ $K_2 \approx 10^6$ $\alpha \approx 10^6$,

$$N/N_0 \approx \alpha/2K_1 \approx 1.$$

Secondary Electron Emission.—Here the mechanism is supposed to be fast electron → soft X-ray → slower electron which escapes. Reabsorption of soft X-ray and re-ejection as slower electrons are almost complete and instantaneous, so that the smaller absorption coefficient of the more penetrating secondary electrons will settle how many of these get out. The values are thus—

$$K_1 \approx 10^6 \; K_2 \approx 10^6 \; \alpha \approx 10^6,$$
$$N/N_0 \approx \alpha/2K_1 \approx 1.$$

According to this scheme a primary electron produces very nearly the same number of soft X-ray quanta as of secondary electrons per cm. of its track, so that the difficulty with Krefft's results does not arise. The small X-ray efficiency is due, not to inefficiency of conversion of electronic into X-ray energy, but abnormal absorbability of the X-rays produced. The efficiency of soft X-ray excitation is of the order 10^{-6} and the efficiency of photo-electric action and of secondary electron emission each of the order unity. These seem to be about the kind of magnitudes required by the experimental results. There may be some simpler scheme than this which unites all these somewhat diversified effects, but I have not been able to think of one.

There is one objection that has to be considered. If there are all these low energy secondary electrons, how is it that the X-ray beam is almost free from radiation having $h\nu <$ about 10 volts. An explanation of this which seems adequate is that at these low speeds the electron collisions appear to be elastic and there is no corresponding conversion into radiation.

A difficulty which has been felt for some time in the soft X-ray field is that no discontinuities have been well authenticated which are characteristic of the material of the photo-electric target. This seems to be explained by the secondary electron experiments, as it appears that no discontinuities are found in these experiments when the targets are not heated to a high temperature. It seems, therefore, likely that the photo-electric inflections have been concealed by gas films or whatever it is that modifies the properties of the surface of a cold metal in a vacuum.

The scheme for secondary emission to which we have been led is not the same as any of the schemes (*a*), (*b*), (*c*) or (*d*) (p. 534), but is closest to (*b*), which was rejected on the ground that the proportion of the radiation in the soft X-ray beam which could pass through fluorite was very low. A good deal of the force of this argument has been removed by the introduction of the hypothesis of *special absorption of the soft X-radiation at the region where it* originates. It does not seem now that (*b*) can be ruled out with certainty, and, in fact, the processes involved may be quite similar to those which occur in the excitation of hard X-rays.

There is a possibility of a misfit by a factor of about a million similar to that here encountered in some other phenomena, involving the interaction of radiation and electrons. In a number of papers starting in 1912,* I put forward the view that chemical reaction, evaporation and thermionic emission, all of which have a close relation to one another, might be due to the integrated photo-chemical or photo-electric effect of the black-body radiation present at the temperature at which the effect occurred. It was easy to show that this hypothesis gives the correct relation between the relative magnitudes of the effects to be expected at different temperatures. At that time there were no data to test the absolute magnitudes. Looking at the matter broadly, there is no very fundamental difference between these effects. Perhaps thermionic emission is the simplest. At any rate, it is the one with which I personally felt most familiar. Accordingly, I undertook the task of ascertaining the requisite photo-electric data with the result† that the thermionic emission was found to be of the order of 10^6 times that to be expected as the integrated effect of the black-body radiation in the case of platinum. This led to the conclusion that photo-electric action could account for but a small part of thermionic emission and that the main thermionic mechanism must be something else. This is not a difficulty on any free electron theory of metallic conduction, but it might be on some others which have been put forward in which there does not seem to be any mechanism other than radiation which will get an electron out of a metal. In any event, the case of thermionic emission does not stand alone. It is not improbable that evaporation and chemical reactions of the type which have high temperature coefficients are actuated by mechanisms which are not fundamentally different from that which actuates thermionic emission. In particular, the case of monomolecular chemical reactions seems to leave no room for any mechanism other than radiation, especially since it has been shown by Tolman[*] in the well-authenticated case of nitrogen pentoxide that activation by collisions is far too slow to account for the measured rate of decomposition. It is also of interest to note that what would seem to be the straightforward way of calculating the rate of this reaction on the hypothesis of a radiation mechanism[†] gives a rate much below that actually found.

---

* O. W. Richardson, 'Phil. Mag.,' vol. 23, p. 615 (1912); vol. 24, p. 574 (1912); vol. 27, p. 477 (1914); vol. 31, p. 149 (1916).

† K. T. Compton and O. W. Richardson, 'Phil. Mag.,' vol. 26, p. 549 (1913); O. W. Richardson and F. J. Rogers, 'Phil. Mag.,' vol. 29, p. 618 (1915); and O. W. Richardson, 'Phil. Mag.,' vol. 31, p. 149 (1916).

These discrepancies have recently been considered at some length by S. C. Roy,[‡] who is of the opinion that the actual mechanism is operated by radiation in these cases, but that there is a fundamental difference in the rate of operation of the mechanism by radiation in thermal equilibrium with its surroundings when compared with the case of radiation acting on a cold system. The efficiency of radiation acting on a cold system is found by using plausible hypotheses to be about a million times less than that of radiation in temperature equilibrium, and, in fact, a fair quantitative account of the observed facts both of the complete photo-electric emission and of monomolecular reactions is given. However, it seems that there is a precisely similar discrepancy when we try to interlink the generation of soft X-rays and of secondary electrons with the photo-electric action of soft X-rays, and in this case we are dealing entirely with cold systems.

We are thus led to enquire again what it is that the systems which go fast have in common, which differentiates them from the systems which go too slow. Evidently it is not that they are in thermal equilibrium with the radiation. It seems to be that they are systems in which the radiation which is absorbed is generated. The phenomena occur in such a way as would happen if the radiation used in firing projectiles such as electrons or the products of chemical reaction were absorbed about one million times as fast in the neighbourhood in which it is generated as would be found to be the case if the same radiation were separated out into a beam and then tested in another place. There is evidence that this holds for all radiations whose frequencies do not exceed those of soft X-rays. The question of the behaviour of waves of higher frequency in this respect seems to be an open one at present.

As the subject matter of this paper is rather intricate, it may be worth while to recapitulate very briefly what appear to be the essential steps in the main argument. The processes involved may be represented by the diagrammatic sketches (fig. 1 and fig. 2).

---

[*] R. C. Tolman, 'Journ. Amer. Chem. Soc.,' vol. 47, p. 1524 (1925).

[†] *Cf.* L. S. Ornstein, 'Faraday Society, Oxford Meeting' (1925).

[‡] S. C. Roy, 'Roy. Soc. Proc.,' A, vol. 110, p. 543 (1925), and vol. 112, p. 599 (1926).

Fig. 1 represents the typical case of soft X-ray excitation (process (1) $e \rightarrow X$) on the left and photo-electric detection of the rays (process (2) $X \rightarrow e$) on the right. Fig. 2 represents the case of secondary electron emission (process (3) $e \rightarrow e$). In each case let N and N_0 be the number respectively of primary and secondary electrons or quanta as the case may be. In Case (3) the experiments show that $N/N_0 \simeq 1$. In Case (1) let $N/N_0 = \xi$ and in Case (2) let $N/N_0 = \eta$. We do not know ξ and η separately, but the experiments show definitely[*] that $\xi\eta \simeq 10^{-6}$. But the discontinuities in process (3) agree with the discontinuities at the anticathode process (1), and not with discontinuities at the photo-electric target in the soft X-ray excitation. They are also comparable fractions of the total effect in each case. This precludes a low photo-electric efficiency ($X \rightarrow e$), which would require the discontinuities in (3) to be correspondingly small as compared with (2) and (1).

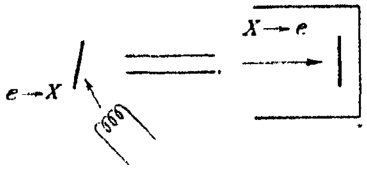


FIG. 1.

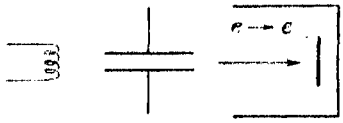


FIG. 2.

We are thus led, using very rough magnitudes, to $X \rightarrow e \simeq 1$, *i.e.*, $\eta \simeq 1$, $\xi \simeq 10^{-6}$. We have seen, equation (5), that unless the absorption coefficients K_1 and K_2 of the primary and secondary beams have entirely different magnitudes, $N/N_0 \simeq \alpha/2K_1$, where α is the number of secondaries generated per unit length of the track of the primary radiation. The values got by direct measurement would lead to K_1 and K_2 each about 10^6 in every case. Adopting a value of this order of magnitude for (2) and (3), we get $\alpha \simeq 10^6$ in each of these cases. If in (1) we retain $K \simeq 10^6$ as called for by direct experiment, we require $\alpha = 1$. But this would mean that in (3) the primary electrons are converted into secondaries about 10^6 times as fast as into X-rays, which would make impossible a roughly proportional correspondence of the soft X-ray and secondary electron inflexions. Hence we are driven to $\alpha \simeq 10^6$ for all cases, $K \simeq 10^6$ for (2) and (3), and $K \simeq 10^{12}$ for (1). This would imply a special absorption of the soft X-rays near their point of origin, as has been pointed out already.

---

[*] Actually $\xi\eta$ is a function of the primary voltage, and it traverses a range from about 10^{-7} to 10^{-5} in the range of voltages applicable to soft X-rays.

The Photo-Electric Properties of some Metals in the Soft X-Ray Region.

By Letitia P. Davies, B.Sc., King's College, London.

(Communicated by O. W. Richardson, F.R.S.—Received March 12, 1928.)

In a paper by Richardson and Robertson[*] dealing with the absolute values of the efficiencies of various substances as soft X-ray producers, the values obtained for carbon, iron, nickel and tungsten differ considerably from those obtained earlier by Richardson and Chalklin.[†] It was suggested that the difference between the two sets of results might be due to the fact that Richardson and Chalklin used a copper photo-electric detector, while Richardson and Robertson used a nickel one. The photo-electric detector consists of a metal disc (the actual detector) surrounded by a cylinder. There is an accelerating voltage (approximately 100 volts) between the disc and cylinder, so that photo-electrons emitted from the disc are drawn to the cylinder. It was thought that the difference in the metals used for the photo-electric disc might account for the difference in the results.

A quartz tube, of similar design to that used by Richardson and Robertson, is used to investigate the suggested cause of difference between the results discussed above. The apparatus is shown in the diagrams below (figs. 1, 2 and 3).

The anticathode A has four sides, at 90°, the four metals forming the targets being iron, cobalt, nickel and copper, of atomic number 26, 27, 28 and 29 respectively. The size of the targets is 2 by 1·4 by 0·1 cm. The anticathode is built up on a nickel framework (fig. 3). The bottom plate L, perforated at the corners, has four upright pins. The four targets (Fe, Co, Ni, Cu), in the order shown in the diagram, slide down between the pins by means of vertical grooves. A section of any one of the targets is shown in fig. 4. A hollow square of nickel fits on to the pins in order to keep the targets tightly in place. Into the bottom plate is screwed a nickel tube G, and a brass rod B passes up through the quartz stopper I into this nickel tube. The latter rests on the pin K, which is perpendicular to the brass rod. The anticathode can be turned about the brass rod by means of an external electro-magnet, acting on the soft iron pieces S, which are screwed into the nickel tube. The tube has four

---

[*] 'Roy. Soc. Proc.,' A, vol. 115, p. 280 (1927).

[†] 'Roy. Soc. Proc.,' A, vol. 110, p. 273 (1926).

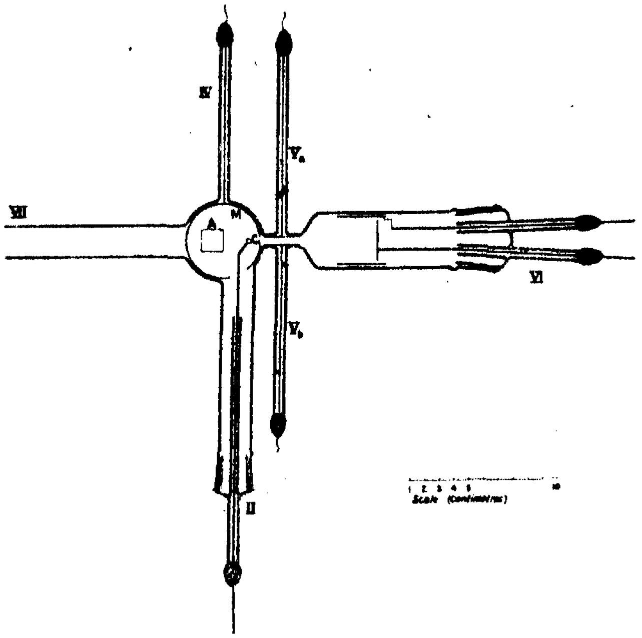
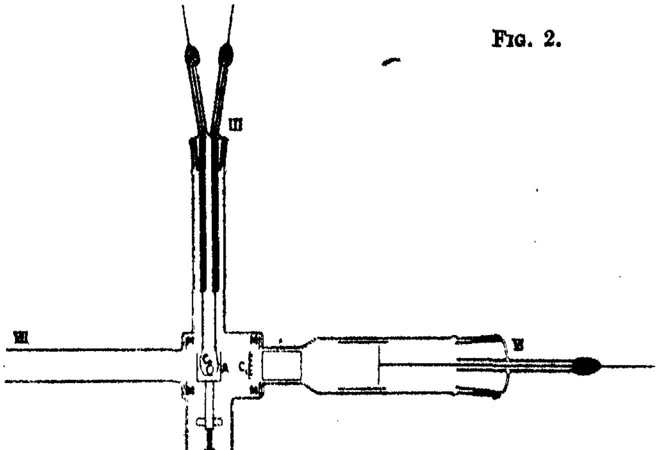


FIG. 1.

FIG. 2.



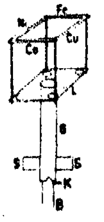


FIG. 3.

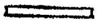


FIG. 4.

notches at equal distances round its circumference. When any one of these notches rests on the pin K, one of the targets is in position for working, *i.e.*, is opposite the filament C_1. As the anticathode is turned so that each notch rests on the pin, each target in turn comes opposite the filament.

The spiral filament C_1 is the source of the electrons producing the soft X-rays. This is of 2-ampere tungsten wire, of length 2 cm., and has 5 turns of diameter 0·3 cm. The ends of the filament are gripped by two copper leads, which pass through the stopper II. The apparatus is provided with a second tungsten filament C_2, in the form of a loop with one turn. This filament (whose leads pass through stopper III) goes down inside the anticathode and is used to bombard the latter in order to remove gas from the targets.

The nickel cylinder M shields the spiral filament and anticathode from the effect of outside charges, and also prevents electrons from the filament or target from reaching the photo-electric detector, being kept at a potential of − 15 volts. A wire from this cylinder is brought out through the tube IV.

The soft X-rays, generated by the anticathode, pass through a slit 0·4 by 2·03 cm. in the nickel cylinder and then between two nickel condenser plates, whose length is 2·65 cm., heights 1·92 and 1·85 cm. respectively and distance apart 0·55 cm. These plates, at + 300 and earth potential respectively, prevent positive ions from reaching the photo-electric detector. The leads from these plates emerge along the tubes V_a and V_b in the diagram. A preliminary experiment showed that under the above arrangement of potentials no positive ions or electrons could pass through to the photo-electric detector.

The rays, after passing through the condenser plates, fall on the photo-electric detector, a disc of diameter 2·7 cm., surrounded by a cylinder of diameter 3·15 cm., and length 3·47 cm. The disc is at a distance 2·82 cm. from the end of the cylinder nearest the condenser plates. The leads from the disc and cylinder respectively are brought out through the stopper VI. The disc lead is connected to one pair of quadrants of an electrometer (the other pair of quadrants being earthed) and the cylinder lead is connected to + 100 volts. This voltage serves to draw the photo-electrons emerging from the disc to the cylinder. The cylinder used in these experiments is of nickel, the metal used for the disc being in turn iron, cobalt, nickel, copper, *i.e.*, the same metals as are used for the anticathode.

The apparatus is evacuated by way of the tube VII. All joints are sealed with hard sealing wax. A Hyvac rotary pump is used to back a mercury diffusion pump; this pumping system quickly evacuates the tube down to a pressure of 10^{-7} mm. Two traps surrounded with liquid air prevent mercury (from the diffusion pump) from reaching the tube, and the trap nearest the apparatus contains charcoal, which absorbs any vapour from tap grease, etc., and reduces the pressure in the tube to the order of 10^{-8} mm. The pressure in the apparatus is measured by a McLeod gauge.

When the apparatus is first put up, the charcoal is baked out so that any gas it has absorbed from the atmosphere is pumped away. When the pressure has fallen to the order of 10^{-5} mm., the quartz tube itself is baked out with a blowpipe and is made red-hot, all sealing-wax joints, of course, being kept cool. In this process usually the pressure rises to about 3 by 10^{-4} mm., but quickly falls again. The anticathode is next bombarded with electrons from the loop filament, the latter carrying a current of 2·3 amperes, 600 volts being applied between anticathode and filament. After about one minute the targets become red-hot and the pressure rises at first to about 5 by 10^{-4} mm. As the bombarding treatment continues the pressure does not rise above 2 by 10^{-5} mm. The bombardment is usually done at first in series of 3 minutes at red-heat; later, when the pressure only rises to 10^{-5} mm., for 10 or 15 minutes at a time. Bombardment is always performed at the beginning of the day, before taking readings, the latter being taken from 6 to 9 p.m. The pressure when readings are taken is always of the order of 10^{-7} mm.

A diagram of the electrical connections is given below (fig. 5).

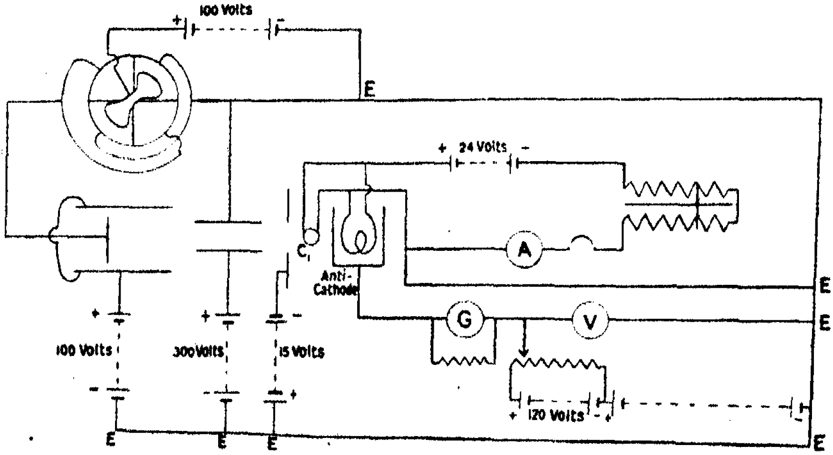


FIG. 5.

The part of the tube containing the photo-electric detector is covered with silver foil, which is earthed, for screening purposes. The quadrant electrometer is of the Dolezalek type; the lead to it is surrounded by an earthed tube, and the electrometer itself is contained in an earthed tin. The rate of positive charging up of the photo-electric disc (due to the loss of photo-electrons) is measured by timing a spot of light reflected from a mirror on the electrometer needle over 20 cm. of a scale roughly 1 metre away.

In a subsidiary experiment this spot is timed also with a condenser of given capacity across the electrometer quadrants, the same current being applied, and from the two readings the capacity of the electrometer can be found. The sensitivity of the instrument in millimetres per volt is obtained directly, and hence the absolute value of the current from the photo-electric disc obtained.

The thermionic current between filament and anticathode is measured by a shunted microammeter, and the voltage across them by a Weston voltmeter.

Curves obtained for the different photo-electric plates are give below in fig. 6.

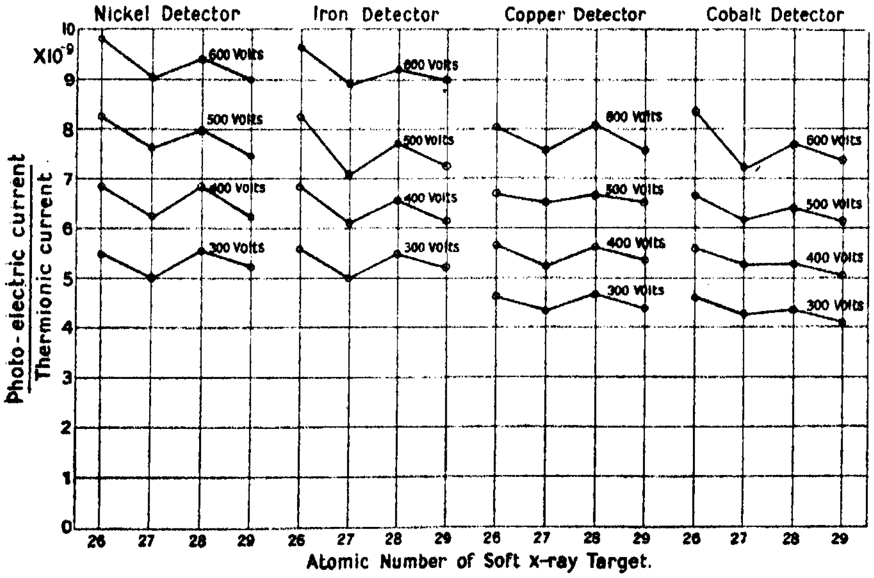


Fig. 6.

The value of the ratio, photo-electric current to the thermionic current (i_p/i_t), is plotted against the atomic number of the target as each target in turn is in the working position. This ratio is usually taken at the four voltages 600, 500, 400 and 300 volts. The points on the curves are the averages of values obtained on not less than two occasions, except in the case of the iron photo-electric detector, on the 600-volt curve. This curve is taken from only one set of readings; on the next occasion on which readings were taken, the high-tension battery would only give 563 volts. The curve obtained at this voltage, however, is of the same shape as that previously obtained for 600 volts.

The absolute values of the ratios (i_p/i_t) vary 2 or 3 per cent. from day to day, but not the ratios of the values for different targets.

The first set of curves are those obtained using a photo-electric detector of nickel. At 600 and 500 volts the iron target gives the greatest value for i_p/i_t, nickel has a value 4 or 5 per cent. less, while cobalt and copper have a value about 8 per cent. lower than that of iron. At 400 and 300 volts the curves are very similar, but the value of i_p/i_t for nickel is approximately the same as that of iron, the values for cobalt and copper being about 8 per cent. less, as before.

The results with an iron photo-electric plate are very similar to those obtained with nickel, the efficiency of iron as a photo-electric detector having approximately the same value as that of nickel. In this case again the nickel target gives a value for i_p/i_t below that of iron but above that of cobalt and copper for the higher voltages; at lower voltages the value for nickel approaches that of iron.

Copper as a photo-electric detector is about 17 per cent. less efficient than iron or nickel. The shape of the curves is also less pronounced in the case of copper, though in other respects the curves are similar to those obtained with nickel and iron. The value of i_p/i_t for the iron and nickel targets is approximately the same. The values of i_p/i_t for cobalt and copper are still the same and are about 6 per cent. less than those for the iron and nickel targets.

The curves obtained using cobalt as photo-electric detector are similar to the curves for copper; the efficiencies of these two metals as photo-electric detectors are approximately the same. In the case of cobalt, at 600 volts the value of i_p/i_t for the iron target is 8 per cent. above that for nickel and 12 per cent. above that for cobalt and copper. At 500 volts the variations are not so pronounced; i_p/i_t for iron is 4 per cent. above that for nickel and 7 per cent. above that for cobalt and copper. At 400 and 300 volts i_p/i_t for cobalt approaches the value for nickel.

Curves are also shown in fig. 7, plotting the ratio i_p/i_t for the different photo-electric plates against the atomic number of the latter, for the four X-ray targets in turn. These are similar in shape to the curves mentioned above, but the variations are much more pronounced. This similarity in shape seems to indicate that the photo-electric efficiency is related to the soft X-ray efficiency of the metals tested.

In finding the absolute values of the various efficiencies, it must be taken into account that only a small fraction of the soft X-rays generated reach the photo-electric detector. This fraction is the ratio of the solid angle subtended by the target face at the photo-electric plate to the total solid angle, *i.e.*, 4π, and this ratio must be evaluated before the absolute values can be obtained.

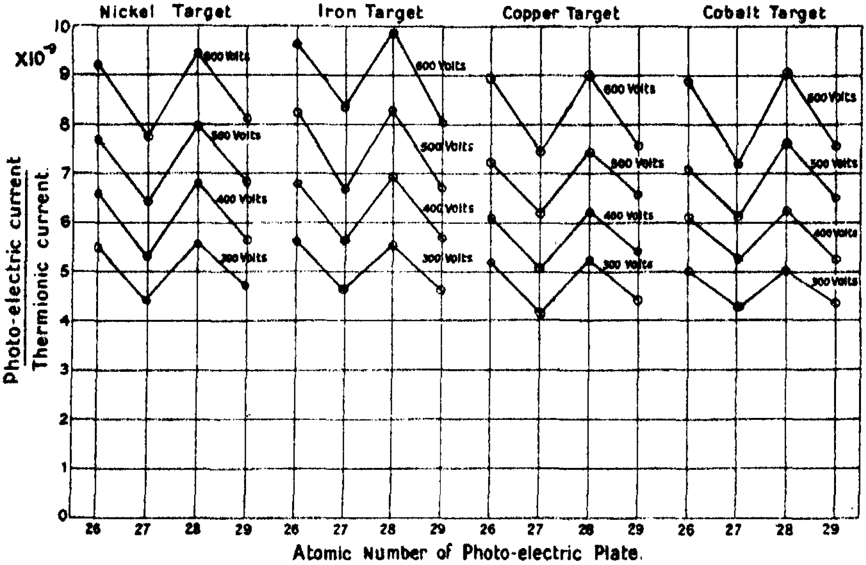


FIG. 7.

The calculation is made from large-scale drawings of the electrodes in the apparatus, in the horizontal and vertical planes. The sines of the angles required are obtained graphically and hence the solid angle found. The value of the solid angle is found to be 0·03047, and thus the absolute value of i_p if all the radiation generated were effective is obtained by multiplying the observed values of i_p by $4\pi/0{\cdot}03047$. In this calculation, however, only that part of the target from which soft X-rays reach the photo-electric plate has been considered. The area of this part is 0·743 of the whole area of the target (this value is obtained from the drawings used for finding the solid angle above); thus the thermionic current to the rest of the target is wasted, *i.e.*, the soft X-rays it produces cannot possibly reach the photo-electric detector. The absolute value of i_p if the whole target were effective is therefore obtained by dividing $4\pi/0{\cdot}03047$ by 0·743.

From some theoretical considerations of Kramers,* mentioned in the paper by Richardson and Chalklin,

$$I = 2\pi \times 3{\cdot}98 \times 10^{-8} \times N_e V e V, \qquad (1)$$

(equation 10 in R and C)

---

* 'Phil. Mag.,' vol. 46, p. 836 (1923).

where N_e is the effective atomic number of the target, e is the electronic charge, V the voltage on the tube, and I is the average radiation energy emitted by the impact of a single electron on the target.

If i_p is the observed photo-electric current,

$$i_p = \alpha n e, \tag{2}$$

where α is the number of electrons emitted by unit incident quantum of radiation, n is the number of quanta of radiation incident on the photo-electric plate in unit time.

It is assumed that the average value of the energy of a quantum is given by $\beta e\mathrm{V}$, where β is a constant and V = voltage on tube. The average energy of n quanta, *i.e.*, the energy of the radiation which gives rise to the observed photo-electric current i_p, is thus

$$\beta n e\mathrm{V} = \frac{\beta i_p}{\alpha}\mathrm{V}. \tag{3}$$

The value of i_p if the whole of the radiation is effective becomes $i_p \times 4\pi/0{\cdot}03047 \times 0{\cdot}743$, so that the energy of the radiation is

$$\frac{4\pi i_p \mathrm{V}\beta}{0{\cdot}03047 \times 0{\cdot}743 \times \alpha}. \tag{4}$$

The number per second of electron impacts generating the radiation is i_t/e, where i_t is the thermionic current to the target; the energy of the radiation for unit electron impact on the target is thus

$$\frac{4\pi i_p e\mathrm{V}\beta}{0{\cdot}03047 \times 0{\cdot}743 \times \alpha i_t}. \tag{5}$$

This is equal to I in equation (1), so that

$$2\pi \times 3{\cdot}98 \times 10^{-8} \times \mathrm{N}_e\mathrm{V}e\mathrm{V} = \frac{4\pi i_p e\mathrm{V}\beta}{0{\cdot}03047 \times 0{\cdot}743 \times \alpha i_t} \tag{6}$$

or

$$\frac{\alpha}{2\beta}\mathrm{N}_e = \frac{1}{\mathrm{V}} \times \frac{i_p}{i_t} \times 1{\cdot}114 \times 10^9. \tag{7}$$

For 1/V = 1, *i.e.*, when V = 300 volts, the following values are obtained for $\frac{\alpha}{2\beta}\mathrm{N}_e$ for the various photo-electric plates:—

Table I.

Element.	Fe	Co	Ni	Cu
Atomic No.	26	27	28	29
$\frac{\alpha}{2\beta}$ N_e (Copper)	5·17	4·86	5·20	4·88
$\frac{\alpha}{2\beta}$ N_e (Nickel)	6·15	5·56	6·20	5·84
$\frac{\alpha}{2\beta}$ N_e (Cobalt)	5·17	4·76	4·87	4·60
$\frac{\alpha}{2\beta}$ N_e (Iron)....	6·23	5·57	6·10	5·72

The values for the nickel photo-electric plate are not very different from those obtained by Richardson and Robertson for the same targets, *i.e.*, Fe, 5·32; Co, 5·03; Ni, 4·85; Cu, 4·10, and the values for the copper plate are very similar, though about 20 per cent. lower. It thus appears that the large difference between the results found by Richardson and Chalklin and Richardson and Robertson cannot be due to the difference in the photo-electric plate.

The values for $i_p/i_t \times 10^9$ at various voltages are given below.

Table II.

Date.	Element.	Atomic No.	600 volts.	500 volts.	400 volts.	300 volts.
		Nickel Photo-Electric Plate.				
May 24th*	Fe	26	9·79	8·38	7·08	5·82
	Ni	28	9·29	7·91	7·06	5·74
	Co	27	8·92	7·57	6·34	5·21
	Cu	29	8·85	7·50	6·36	5·40
	Fe	26	9·86	8·40	7·18	5·89
May 26th*	Fe	26	9·88	8·05	6·60	5·18
	Ni	28	9·55	8·00	6·59	5·38
	Co	27	9·18	7·69	6·18	4·76
	Cu	29	9·20	7·44	6·07	5·07
	Fe	26	9·82	8·19	6·60	5·19
May 31st	Fe	26	10·0	8·31	6·71	—
	Ni	28	9·63	7·77	6·25	—
	Co	27	9·18	7·92	5·91	—
	Cu	29	9·12	7·43	6·05	—
	Fe	26	9·97	8·37	—	—

Table II—(continued).

Date.	Element.	Atomic No.	600 volts.	500 volts.	400 volts.	300 volts.
		Iron Photo-Electric Plate.				
June 10th	Fe	26	9·6*	7·98	—	—
	Ni	28	9·17	7·77	—	—
	Co	27	8·87	7·10	—	—
	Cu	29	8·96	7·34	—	—
	Fe	26	—	7·83	—	—
			(563 volts)			
June 21st	Fe	26	8·71	7·98	6·84	5·97
	Ni	28	8·88	7·95	6·88	5·95
	Co	27	8·23	7·39	6·68	5·75
	Cu	29	8·35	7·61	6·68	5·63
	Fe	26	9·02	8·06	6·91	6·02
			(600 volts)			
June 22nd*	Fe	26	—	8·13	6·84	5·41
	Ni	28	—	7·73	6·57	5·46
	Co	27	—	7·13	6·11	4·89
	Cu	29	—	7·26	6·16	5·20
	Fe	26	—	8·35	6·60	5·40
June 23rd*	Fe	26	—	8·31	6·94	5·81
	Ni	28	—	7·64	6·54	5·47
	Co	27	—	7·00	6·11	5·11
	Cu	29	—	7·20	6·08	5·23
	Fe	26	—	8·30	6·93	5·73
		Copper Photo-Electric Plate.				
July 13th*	Fe	26	8·13	6·89	5·80	4·59
	Ni	28	8·03	6·83	5·66	4·70
	Co	27	7·49	6·65	5·24	4·71
	Cu	29	7·60	6·78	5·54	4·52
	Fe	26	7·99	6·92	5·74	4·72
July 15th*	Fe	26	8·05	6·74	5·54	4·59
	Ni	28	8·14	6·75	5·57	4·64
	Co	27	7·67	6·34	5·19	4·31
	Cu	29	7·51	6·27	5·20	4·24
	Fe	26	8·13	6·16	5·56	4·67
		Cobalt Photo-Electric Plate.				
July 26th*	Fe	26	8·38	7·10	6·00	5·00
	Ni	28	8·04	6·83	5·73	4·83
	Co	27	7·35	6·40	5·66	4·65
	Cu	29	7·71	6·53	5·57	4·53
	Fe	26	—	—	5·86	—
July 27th*	Fe	26	8·31	6·48	5·21	4·28
	Ni	28	7·33	5·96	4·81	3·90
	Co	27	7·06	5·93	4·85	3·88
	Cu	29	7·06	5·80	4·52	3·72
	Fe	26	7·70	6·36	5·28	—

The results marked * are those from which curves are plotted, the average value for each point being taken.

In conclusion, I am glad to have this opportunity to thank Prof. Richardson for suggesting this problem and for supervising the experiments.

Condensation Phenomena at Different Temperatures.

By C. F. Powell, B.A., Sidney Sussex College, Cambridge.

(Communicated by C. T. R. Wilson, F.R.S.—Received March 17, 1928.)

§ 1. *Introduction.*

The results of the early experiments of C. T. R. Wilson[*] on the condensation of water vapour have been of importance to both pure and applied physics. They led to the development by Wilson himself of the cloud method of investigating atomic phenomena, and, in the hands of H. L. Callendar[†] and others, they were applied to the thermodynamic theory of the steam turbine, especially in connection with the discharge of steam by nozzles.

Wilson's results, together with those of the earlier investigations of Coulier, Aitken, etc., may be summarised as follows: Suppose we have a glass vessel, the volume of which can be altered at will, containing, initially, a volume v_1 of air taken from the atmosphere of a room and saturated with water vapour at room temperature. If the air is allowed to expand suddenly to a volume v_2, so that the ratio of the final volume to the initial volume, v_2/v_1, is slightly greater than one, a dense cloud of drops is formed in the vessel. On restoring the gas to its original volume, a second expansion like the first results in the formation of a similar cloud, but if this process is repeated a stage is reached when the clouds become thinner, the drops being correspondingly larger, and soon the result of an expansion is not a dense cloud of small drops, but a rain of quickly falling large ones. Further small expansions now result in no condensation throughout the gas.

If the gas is now subjected to expansions of gradually increasing size, the following phenomena may be observed: As long as the quantity v_2/v_1, which may be called the expansion ratio, is less than $1 \cdot 25$, no condensation results, except on the walls of the containing vessel, but at this limit a few drops appear as a result of the expansion. With larger expansions the number of drops increases slowly until a second limit is reached, when the expansion ratio $= 1 \cdot 38$. Beyond this limit water is condensed, not in the form of a thin rain of heavy drops, but as a persistent cloud of small ones. The number of drops increases rapidly with v_2/v_1, each becoming correspondingly smaller, since the amount of water condensing at each expansion changes but little, and diffraction colours begin to border the image of a source of light viewed through the cloud. When the size of the drops becomes comparable with the wave-length of light, the diffraction system expands, and its centre, originally white, becomes coloured.

---

[*] 'Phil. Trans.,' vol. 189, p. 265 (1897).

[†] 'Proc. Inst. Mech. Eng.' (1915).

The explanation of these experiments is now well known. The clouds first formed are produced by condensation on the dust particles present in the atmosphere, which eventually are removed with the water that condenses on them. The nuclei on which water condenses when $v_2/v_1 = 1 \cdot 25$ are charged ions, continually being renewed throughout the gas, whilst those which become effective when $v_2/v_1 = 1 \cdot 38$ are uncharged, and have been identified with the associated molecules present in water vapour.

Assuming (*a*) that the volume of the mixture of air and vapour changes adiabatically, and (*b*) that no appreciable condensation takes place until the expansion has been completed, it is easy to calculate the absolute temperature of the mixture at the end of the expansion from the equation

$$\left(\frac{v_2}{v_1}\right)^{\gamma-1} = \frac{T_1}{T_2}. \tag{1}$$

In his experiments Wilson, using apparatus which differed widely both in capacity and speed of expansion, found the same values for the expansion ratio required to produce the condensation on a given kind of nucleus, and he was therefore justified in making the assumptions indicated above, and in applying equation (1) to his results.

In this way he found that with an initial temperature of 25° C., an expansion for which $v_2/v_1 = 1 \cdot 25$ produced a lowering of about 25° C. in the temperature of the mixture. Defining the supersaturation as the ratio of the actual density of vapour to the density of saturated vapour at the same temperature, he deduced that a fourfold supersaturation is necessary to produce condensation on ions at 0° C. Similarly, he found that the supersaturation necessary to produce the cloud-like condensation which appears when $v_2/v_1 = 1 \cdot 38$ is eightfold at $-$ 16° C.

In the experiments described in this paper, an attempt has been made to investigate the phenomena more fully by making experiments at different temperatures and with varying pressures of air. From the results it has been possible to discuss the application of the cloud method of investigating atomic phenomena in a wider range of conditions than has hitherto been used. The supersaturation required to produce the cloud-like condensation at different temperatures has been found, and the significance of the results in connection with the theory of the steam turbine explained.

To avoid circumlocution, the expression "expansion at the cloud limit" will be used instead of "the expansion ratio required to produce condensation on the nuclei which become effective when the cloud limit is reached." An expression in which "supersaturation" takes the place of "condensation" will, *mutatis mutandis*, be used in a similar way.

§ 2. *The Apparatus.*

The final form of the expansion chamber used in these experiments is shown diagrammatically in fig. 1. It is made entirely of glass, and consists of a cylindrical piston A, 2 cm. in diameter and 3 cm. long, sliding in an outer tube BD bearing a constriction C, into which the conical end of the piston is ground. The instrument is made from tubing of 2-mm. wall thickness, carefully chosen to be circular in section and very slightly conical. The piston is ground first on the lathe until it will just enter the wider end of the outer tube, and then into the latter by hand. For this purpose the flat end of the piston is attached to a mandrel by means of shellac. The success of the instrument depends upon the skill with which the constriction and the conical part of the piston have been made, for it is very difficult to locate the abrasive only round the conical portions when these are being ground together. If they are both symmetrical and of the same angle, the time taken in grinding them together is small, and but little glass is removed from the walls; otherwise this time is long, and the clearance between the walls and the piston becomes too great, so that when the piston is made to fly into the constriction, air is entrained between the piston and the outer walls, and splashing in the chamber makes it difficult to estimate the effects of the expansion. By using the finest grades of emery powder the joint at the constriction can be made air-tight, when lubricated with water, even with a pressure difference of two atmospheres between the two sides of the piston.

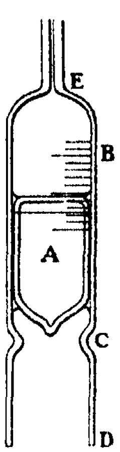


FIG. 1.

For experiments in which large numbers of expansions have to be made, it is inconvenient to measure the expansion ratio by observations on the pressure. A scale is therefore etched on the inside of the expansion chamber and a line round the piston, all the markings lying in planes perpendicular to the common axis of these two. The instrument was calibrated as described in the next paragraph, so that the expansion ratio corresponding to a given initial position of the piston is known. After the scale has been etched, one end of the outer tube is cut off to within 2 cm. of the constriction and ground flat, and the other is drawn off and joined at E to a capillary, as shown in the diagram.

Apart from the method of measuring the expansion ratio, the instrument is very similar to one used by Wilson. The advantages of the all-glass apparatus seem to be threefold :—

(*a*) All risk of contamination can be avoided;
(*b*) The expansion suffered by the vapour is capable of measurement to a high degree of accuracy; and
(*c*) The end of the expansion is very sharp.

These advantages make it eminently suitable for work of precision. Some preliminary experiments were made with a chamber in which a piston fell on to rubber, but with a yielding stop of this kind it is difficult to ascertain the maximum expansion suffered by the vapour.

The water bath, shown in fig. 2, consists of a vertical glass cylinder A, 10 cm. in diameter, fitting into flanges turned in the brass discs B, C, which form the top and bottom. Rubber rings in the flanges ensure that the joint between the glass and the brass is water-tight when the discs are drawn together by three rods, carrying lock-nuts, which pass through the bath.

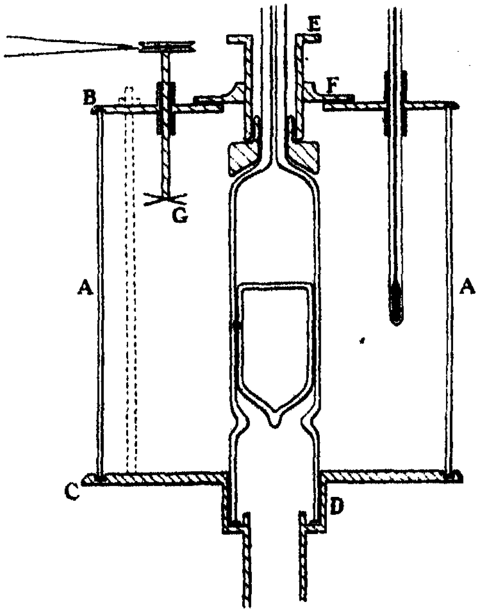


Fig. 2.

The base of the expansion chamber is pressed against a rubber ring in a guide

D, at the base of the bath, by a tube E, which screws into a piece F. When the three screws which normally attach F to the top of the bath are withdrawn, the expansion chamber can be removed for cleaning purposes without dismantling the bath.

The water in the bath is heated electrically by a coil carried in a tube passing round the base of the bath, and it is stirred vigorously by a propeller G, driven by a small electric motor. The tube carrying the coil is filled with glycerine to ensure good thermal contact between the coil and the water in the bath. With this arrangement it is easy to maintain the temperature constant to 0·1° C. by adjusting a variable resistance in series with the heating coil.

An expansion is made by suddenly reducing the pressure beneath the piston by pulling away a rubber bung from the end of a tube which leads to a Winchester bottle, exhausted by the continuous action of a water pump. The pressure beneath the piston is thus rapidly reduced, and the piston falls into the constriction. The method is too familiar to need more detailed description.

The complete apparatus is shown in fig. 3. For lubrication purposes water is run into the chamber after each expansion through the tap G from the vessel J, where it is stored under pressure. In order to avoid contamination from the rubber bung closing the end of J, it is always covered with mercury to a depth of 1 cm. With this precaution, and especially if the piston is left at the top of the chamber when the apparatus is not in use, it is some weeks before the chamber needs cleaning. When the chamber is clean the water runs over the surface like a thick oil, and contamination is immediately apparent because of the break-up of the surface into ill-defined patches. Freshly boiled distilled water is always used in these experiments. On reducing the pressure in J by lowering the reservoir K, the tap H being open, air can be withdrawn from the chamber by opening the tap G. Conversely, air can be introduced by compressing it in J and opening G.

Communication between the Winchester bottle and the water pump can be cut off by the tap O. The pressure in the bottle, which is always as low as possible unless otherwise stated, is shown on the manometer N.

The other parts of the apparatus are designed to enable the pressure beneath the piston to be both varied and measured, accurately and quickly, over a range from 0–2 atmospheres. Pressures above that of the atmosphere are produced by allowing mercury to flow from the large reservoir L, through the tap S, to the vessel E; those below it, by opening the tap A which communicates with the exhausted Winchester. B controls a capillary leak to *vacuo* for fine adjustment, whilst C and D are fast and slow leaks to air the atmosphere respectively. The pressure in the apparatus, which differs from that in the expansion chamber only by the difference required to support the weight of the piston, is measured either by the open-end manometer P, or by the closed-end manometer M, according to its magnitude.

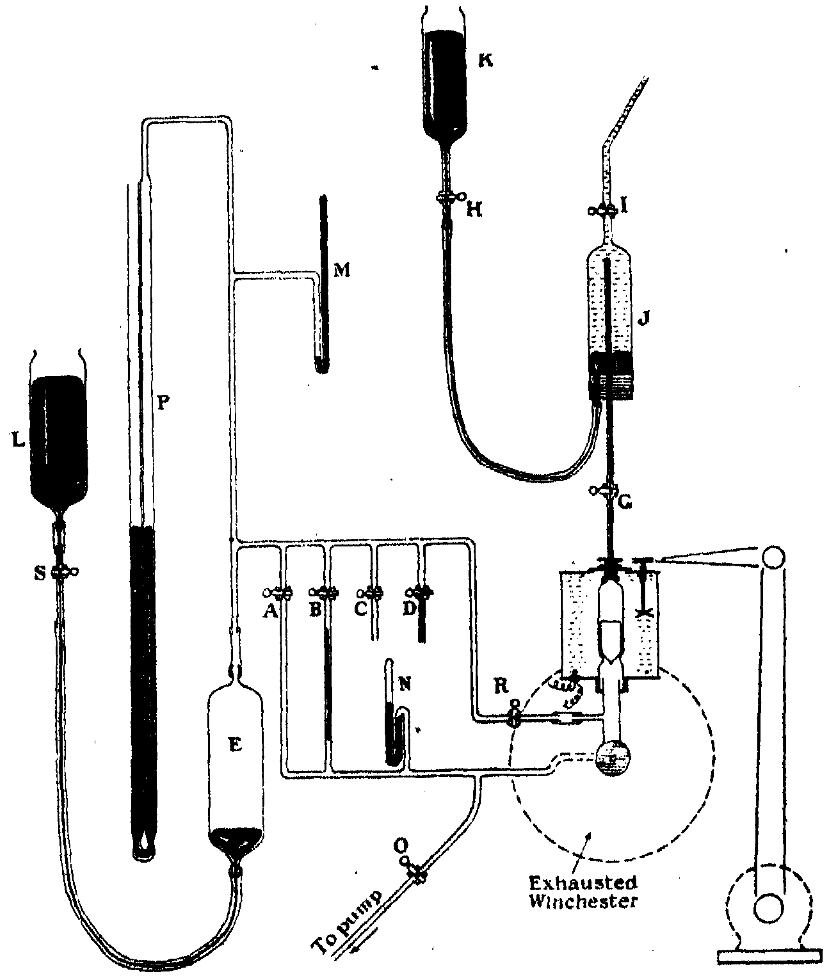


Fig. 3.

In making an expansion the pressure is adjusted until the piston comes to the required initial position, the tap R is closed, and the bung, which normally closes the communication between the exhausted Winchester and the space beneath the piston, is pulled off. The piston falls into the constriction and the results of the expansion are noted.

The illumination is varied according to the nature of the condensate. For investigating the few drops produced at the ion limit a narrow beam of light from the filament of an electric lamp is concentrated by a lens into the chamber, whilst for the examination of the coloured fogs the lens is either removed or made to bring the light to a focus before it reaches the chamber. The contents of the chamber are examined by looking into it along a line which is just out of the direct path of the beam.

§ 3. *Measurement of the Expansion Ratio.*

The position of the piston is defined by the position of the mark on it relative to the scale on the walls of the chamber. If the mark coincides with the xth division of the scale, the piston may be said to be "at" x.

The chamber is calibrated at room temperature. When the piston is at x let the pressure beneath it be P_1. If the pressure is gradually reduced the piston falls and, the motion being slow, the air and vapour change in volume isothermally. When the piston reaches its invariable expanded position, y, let the pressure beneath it be P_2. If p_v, σ represent the vapour pressure of the water in the chamber and the pressure required to support the weight of the piston respectively, then, applying Boyle's Law to the air in the chamber,

$$\frac{v_2}{v_1} = \frac{P_1 - p_v - \sigma}{P_2 - p_v - \sigma},$$

where v_2, v_1 are the volumes of the mixture of air and vapour when the piston is at y, x respectively. Corresponding to different values of x, we can deduce the values of v_2/v_1 and a curve showing the relation between v_2/v_1 and x can be drawn. In order that the expansion ratio may be determined accurately, it is necessary that the mark on the piston shall always be parallel to those on the walls. The markings must, therefore, be perpendicular to the axis of the surface on which they are etched, and the clearance between the walls and the piston must be small, so that their axes may always coincide. These conditions being fulfilled, the accuracy of the calibration is limited by the variations in the amount of lubricating water above the piston. When clear of the constriction, the piston floats up through any excess water that may be run in above it, and at equilibrium there is a certain distribution of water above and below, the excess draining away. This distribution remains the same until the piston approaches the constriction, when water tends to accumulate below it, the amount above decreasing. In measuring P_2 it is necessary, therefore, to adjust the amount of water above the piston in order to avoid a slight systematic error in the calibration curve.

Throughout these experiments the position of the piston has been estimated by eye, and the consistency of the results indicates that v_2/v_1 has been measured to within 0·5 per cent. The pressure required to support the piston is very small, and has been neglected.

§ 4. *The Supersaturation produced by an Expansion.*

The final temperature of a mixture of two perfect gases, expanded adiabatically, can readily be found if we know γ_1, γ_2, the ratio of the specific heat at constant pressure to that at constant volume for each. If the partial pressures for the two components are p_1, p_2 respectively, the total pressure being P, then it may be shown that γ for the mixture is given by the formula

$$\gamma = 1 \Big/ \left[\frac{1}{\gamma_1 - 1} \cdot \frac{p_1}{\mathrm{P}} + \frac{1}{\gamma_2 - 1} \cdot \frac{p_2}{\mathrm{P}}\right].$$

It has been shown that in the adiabatic expansion of dry steam, the relation between the pressure and volume is given to a high degree of accuracy by the formula $p\,(\mathrm{V} - b)^{1\cdot3} =$ constant. b is sufficiently small to be neglected, and the equation becomes $p\mathrm{V}^{1\cdot3} =$ constant. Regarding water vapour as a perfect gas, γ for a mixture of it with air will be given by the equation

$$\gamma - 1 = 1 \Big/ \left(2{\cdot}50\frac{p_a}{\mathrm{P}} + 3{\cdot}33\frac{p_v}{\mathrm{P}}\right), \tag{2}$$

where P is the total pressure of the mixture and p_a, p_v are the partial pressures of the air and vapour respectively. This formula is accurate when $\frac{p_v}{\mathrm{P}} = 0$ or 1, but will lead to a systematic error at intermediate values. This error will, however, be slight, and, as will be seen later, is of no importance to the main theme of the paper.

It is convenient to draw a curve from equation (2) showing the value of γ for mixtures of different compositions. From it the supersaturation produced when air, saturated with water vapour at a given temperature, is expanded may be found. For suppose the magnitude of the expansion to be v_2/v_1 and the initial temperature T_1° A., it is first necessary to find the initial composition of the mixture, and it is inconvenient to do this by determining the initial pressure before each expansion. Let the pressure beneath the piston be slowly reduced. When the piston just reaches the expanded position, let P_2 be the total pressure of air and vapour in the chamber. The vapour pressure of water at T_1° A. being p_v, the partial pressure of the air $= \mathrm{P}_2 - p_v$. If the pressure beneath the piston is now increased it rises. When the piston reaches the initial position corresponding to the expansion ratio v_2/v_1, the partial pressure of the air will have become $v_2(P_2 - p)/v_1$ and the total pressure

$$v_2(P_2 - p_v)/v_1 + p_v.$$

The quantity p_v/P of equation (2) will therefore equal

$$\frac{p}{\frac{v_2}{v_1}(P_2 - p_v) + p_v}.$$

The γ for a mixture of this composition having been determined, the final temperature T_2, corresponding to the expansion considered, may be found from equation (1). For a given quantity of air in the chamber, P_2 can be found and the initial composition need only be calculated for those expansions which produce significant results.

If the density of saturated water vapour at T_1° A., T_2° A. is ρ_1, ρ_2 respectively, then the density of water vapour at the completion of the expansion equals $\rho_1 . v_1/v_2 = \rho_a$ and the supersaturation S, defined as the ratio of the actual density of the vapour ρ_a to the density of saturated vapour at the same temperature, ρ_s equals $\rho_1 v_1/\rho_2 v_2$.

At room temperature and with air pressures greater than that of the atmosphere, such as Wilson used in his experiments, the quantity p_v/P is of the order of 0·01, and the γ for the mixture is indistinguishable from that for air. In these experiments, however, this quantity has sometimes attained the value 0·7.

§ 5. *Method of Experiment.*

After the expansion chamber has been cleaned by washing with solutions of caustic soda and nitric acid, followed by distilled water, an experiment is made as follows: The resistance in series with the heating coil is adjusted to maintain the bath at the desired initial temperature, and air is introduced into the chamber, the quantity depending on the relative proportions of air and vapour required. At a given initial temperature it is usual to start with the highest air pressure and then to make successive experiments with decreasing quantities.

The quantity of air having been adjusted, a few small expansions are made to remove dust particles, and then a series of gradually increasing ones to determine the expansion ratio required just to produce condensation on ions. A convenient number of these is produced by a γ-ray source placed in a tube in the bath near the expansion chamber. The source is next removed and the cloud limit determined. When a cloud has been formed some of the nuclei on which water condenses persist for some time, and the usual precaution of removing these by a series of small expansions is always taken.

The colour phenomena will not be discussed, but it is sufficient to notice here that a given colour in the centre of the diffraction pattern corresponds to a given size of drop. When the cloud limit has been determined the expansion ratio required to produce a given tint in the centre of the pattern is found. The particular colour used is a delicate mauve, described by Wilson as the sensitive tint.

The following table is an example of the methods of observation :—

November 1, 1927.

Initial Temperature = 50° C. Vapour Pressure at 50° C. = 92·3 mm.

With the piston in the expanded position :—

Total pressure in the chamber = 684 mm.

Partial pressure of air = 592 mm.

Initial position of piston.	Effect of expansion.	v_2/v_1.	Initial pressure of air.	p_v/P.
	γ ray source in bath.			
2·0	No drops			
3·0	Heavy rain			
2·5	No drops			
2·6	,,			
2·7	,,			
2·8	Several drops	1·240	733 mm.	0·113
	γ ray source removed.			
2·9	No drops			
3·9	Rain			
4·2	Cloud			
4·0	Rain			
4·1	Heavy rain			
4·05	Estimated cloud limit	1·310	775 mm.	0·106
5·0	Brilliant green centre			
5·4	Yellow centre			
5·2	Sensitive tint	1·383	819 mm.	0·101

These results are indicated in fig. 4, thus : ⊕ .

The three limits having been determined, a little air is removed from the chamber and the operations described above are repeated.

There is no marked change in the phenomena as the amount of air is reduced, but the size of the drops condensing on ions becomes much smaller than at high pressures. The light scattered by the drops is correspondingly diminished, and it becomes necessary to improve the illumination. Similarly, the clouds become thinner and the colour phenomena much less brilliant than at high pressures.

§ 6. *Results at 50° C.*

As an example of the results obtained, fig. 4 shows the expansion ratio required to produce condensation at

(*a*) the ion limit,
(*b*) the cloud limit, and
(*c*) the sensitive tint,

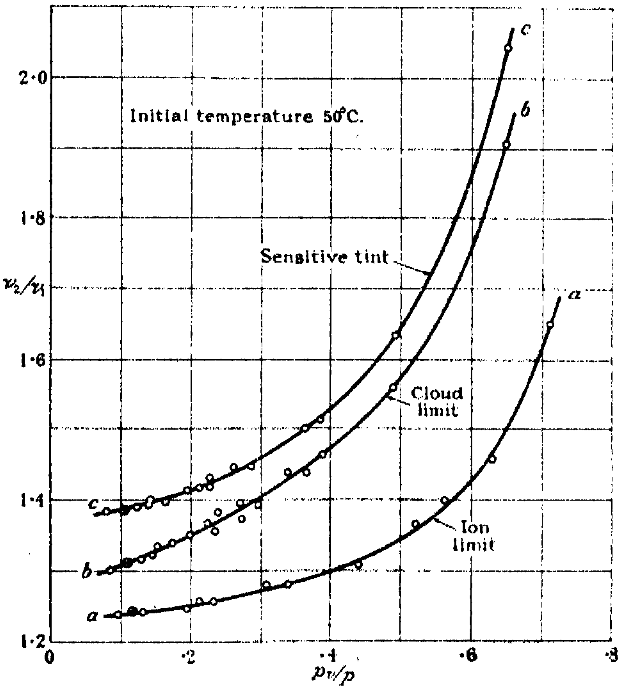


FIG. 4.

as the proportions of air and vapour are varied, the initial temperature being 50° C.

In these experiments the significant quantities are those defining the state of the mixture as the piston comes to rest. These are :—

(1) The final temperature of the mixture ;
(2) The final density of the vapour ; and
(3) The final density of the air,

and they can be calculated for any given expansion, assuming that it is adiabatic and that no appreciable condensation takes place until it is completed.

Thus, in the example given in the preceding table, when $p_v/P = 0\cdot106$, the expansion ratio at the cloud limit $= 1\cdot310$, and γ for such a mixture $= 1\cdot386$. The final temperature T_2, given by the equation $323\cdot1/T_2 = (1\cdot310)^{0\cdot386}$, equals, therefore, 291·1° A. (= 18° C.).

Now the density of saturated water vapour at 50° C. is 0·0838 gram/litre, and at 18° C. 0·0162 gram/litre. Thus the supersaturation at 18° C. at the cloud limit is $8\cdot38/(1\cdot310 \times 1\cdot62) = 3\cdot95$. The density of the air at the end of the expansion may be shown to equal 0·844 gram/litre.

Applying similar assumptions and similar methods of calculation to the other points in fig. 4, we can determine the corresponding values of the supersaturation S and the final temperature, and in fig. 5 the supersaturation at the cloud and ion limits so determined is plotted against p_v/P.

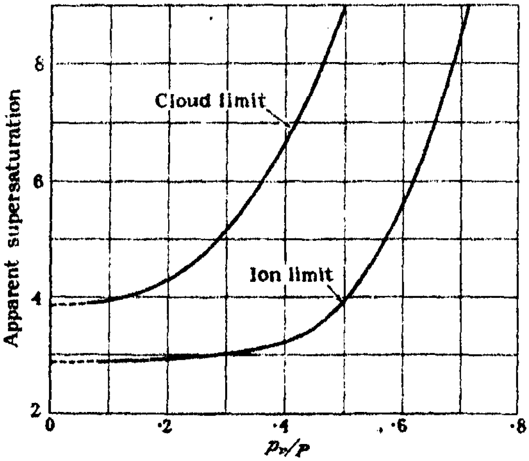


Fig. 5.

Now Wilson showed that the nature of the gas mixed with the vapour had no effect on the supersaturation required to produce condensation on a given kind of nucleus, and he also found that over a small range of temperatures in the neighbourhood of 20° C. the expansion ratio was independent of the initial temperature. We might well conclude, therefore, that the presence of air is without effect on the supersaturation at a given limit, and we should expect this supersaturation to change but slowly with temperature. If the supersaturation at a given limit—*e.g.*, the cloud limit—were independent of the temperature and of the presence of air, the points showing the relation between S at that limit and p_v/P would lie on a straight line parallel to the axis $S = 0$; if the supersaturation changed with the final temperature, but were independent of the presence of air, the line would be slightly curved owing to the variation of γ for mixtures of different composition.

From fig. 5 it will be seen that at high pressures for which p_v/P is small the change of S, as calculated from experiments on both the cloud and ion limits, is slight, but that as p_v/P increases the variation soon becomes very great. Two explanations of this effect are possible: firstly, that the supersaturation at a given limit is profoundly affected by the presence of a foreign gas, an explanation which, especially in the case of the ion limit, is very unlikely; secondly, and more probably, that the assumptions made in calculating the supersaturation are at fault. Since the speed of expansion is much reduced when the total pressure is small, some experiments were made in which the pressure in the Winchester bottle was varied. It was then found that when the total pressure was small so that p_v/P was comparatively large, the result of an expansion of given magnitude varied with the pressure in the reservoir, and, therefore, with the speed of expansion. Evidently the mixture was not suffering an adiabatic expansion of the magnitude deduced from the calibration curve. The view here taken of the phenomena is as follows :—

Before the piston commences to fall the walls of the chamber are covered with a thin film of water, and the air throughout the chamber is saturated. As the piston moves the vapour expands, the air becomes unsaturated with respect to the water film, and this begins to evaporate. The amount of water evaporating during the expansion will depend on

(*a*) the speed of the piston, and
(*b*) the rate of diffusion of water vapour through the air.

At high air pressures the rate of diffusion, which varies inversely as the air pressure, and the time of fall of the piston are both small, and but little water evaporates. Thus Wilson found his results independent of the size and speed of his expansion chamber when these were varied within wide limits. As the partial pressure of the air is reduced, however, the amount of evaporation during the expansion increases rapidly. The density of an element of the gaseous mixture in the middle of the expansion chamber does not change in the ratio v_2/v_1, and the lower the air pressure the greater the departure of the actual change in density from that deduced from the change in volume. Eventually a stage may be reached when, during the expansion, the total pressure in the chamber becomes less than the vapour pressure of the water on the walls. The evaporation will then be controlled not by the comparatively slow process of diffusion, but by the rate at which heat can be communicated to the surface of the liquid, by conduction from the interior of the film, to supply the latent heat of evaporation.

For the apparatus used in these experiments, rough estimates show that if pure steam at 100° C. were expanded instantaneously to twice its original volume, the time taken for enough water to evaporate to restore the pressure to its original value would be of the order of 10^{-4} secs. It seems probable, therefore, that supersaturation sufficient to result in condensation on ions cannot be produced in pure vapours in vessels of ordinary size by the methods used in these experiments. Hirn in steam, and Cazin in other vapours, have produced condensation by the expansion method when demonstrating the difference between vapours having positive and negative specific heats, but the vapour in their experiments was only just saturated and no excess liquid was present. It seems probable, too, that many large nuclei were present, so that the slightest supersaturation would be sufficient to produce condensation. In their experiments any departure from the adiabatic condition was due only to the communication of heat to the vapour by conduction from the walls.

It will be seen that the actual form of the curves in fig. 5 supports the theory just given. At the ion limit the calculated supersaturation changes little with the composition of the mixture until p_v/P reaches the value 0·4, when it increases very rapidly. For the cloud limit the rapid increase occurs at a lower value of p_v/P, and the effect of a change of p_v/P is in general more marked, since the time of expansion and the rate of diffusion for a mixture of given initial composition is greater than in the case of the ion limit.

It may be noticed here that if the speed of expansion is maintained constant, the discrepancy between the actual and calculated values of the density change produced in the gas in the middle of the chamber, due to evaporation of the surface film, will be less in the case of large than of small vessels, since the ratio of the volume of the vessel to the surface area of its walls is proportional to the first power of the linear dimensions for vessels similar in form.

In short, we may conclude that the calculated supersaturation will be equal to that actually produced in the vapour only if the amount of water evaporating from the walls is small. The curves in fig. 5 indicate that this condition is fulfilled only when the pressure of the foreign gas, in these experiments air, is large compared with the vapour pressure.

§ 7. *The Quantity of Water condensed by an Expansion.*

It is interesting to apply to the results of these experiments Wilson's method of determining the amount of water condensing as a result of an expansion. Let ρ_1, θ_1, be respectively the density and temperature of the vapour in the chamber at the end of the expansion before condensation starts, and ρ_2, θ_2 these quantities when the transitory equilibrium, which persists until the influx of heat through the walls becomes appreciable, has been established. Then, if L is the latent heat liberated when one gram of vapour condenses, and C the capacity for heat of each cubic centimetre of the mixture, it is easily shown that $L(\rho_1 - \rho_2) \doteqdot C(\theta_2 - \theta_1)$. The magnitude of C is largely determined by the amount of air present, and for given conditions the quantity of water condensing per cubic centimetre, $\rho_1 - \rho_2$, may be found by trial. On consideration it becomes evident that $\rho_1 - \rho_2$ will decrease with C, and that here is the explanation of the decrease in the size of the drops at the ion limit, with the decrease in the partial pressure of the air in the expanding mixtures.

Similar considerations enable us to explain what at first sight seems an anomaly in the results shown in fig. 4. It will be seen that the slope of the line showing the expansion ratio at the sensitive tint is less than that for the cloud limit when p_v/P is small. This appears to indicate that the increase in the expansion ratio, attributed largely to evaporation from the water film, is less for large than for small expansions. The explanation lies in the fact that at the cloud limit we are determining the expansion ratio required to produce a given number of drops, and at the sensitive tint that for a given size of drop. As p_v/P increases a given supersaturation results in a decreasing amount of condensation, and fewer nuclei are required to give drops of a given size. The clouds therefore become less brilliant, and supersaturation at the sensitive tint approaches that at the cloud limit.

Briefly, the importance of the air or other foreign gas in the expansion method is twofold:—

(*a*) It prevents the evaporation of water vapour from the walls, thus enabling the mixture in the chamber to expand and to become supersaturated; and

(*b*) By increasing the amount of water distributed among a given number of nuclei, it assists in making the drops visible.

We may conclude that for experiments at low pressures the best conditions for producing supersaturation would be :—

(*a*) To introduce into the expansion chamber only sufficient liquid to make the gas just saturated initially ; and

(*b*) To produce the expansion by means of a mechanical device.

§ 8. *The Supersaturation at the Cloud Limit.*

The results of experiments on the cloud limit at temperatures 18° C., 35° C. and 77° C. are plotted together with those for 50° C. in fig. 6, the dotted portions representing extrapolations to the axis $p_v/P = 0$. The results for these temperatures are very similar to those for 50° C., which have already been discussed, and the general form of the curves calls for no special comment. In what follows it will be assumed that the supersaturation at the cloud limit is independent of the presence of air. It has been shown that this hypothesis is compatible both with these experiments and with those of Wilson, and the deductions from it agree so well with experiments on the flow of steam through nozzles, which will be discussed in another place, that there is little reason to doubt its truth.

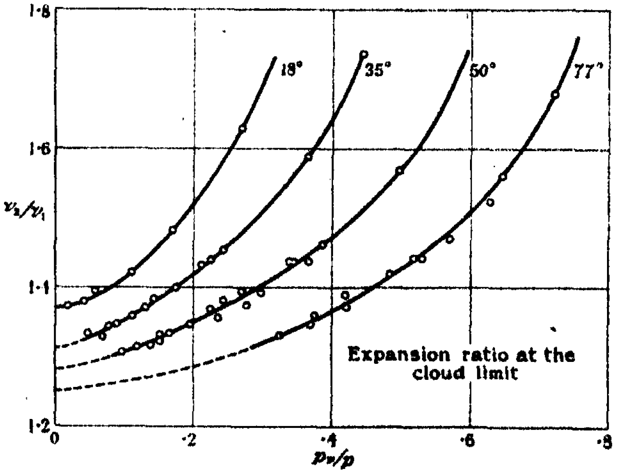


Fig. 6.

The supersaturation at the cloud limit can best be determined from experiments at high air densities, where the expansion is adiabatic and its magnitude deducible from the calibration curve. For convenience, the points where the curves in fig. 6 cut the ordinate $p_v/P = 0$ have been used. With initial temperatures of 50° C. and under, the results obtained in this way are almost identical with those obtained from the actual experiments at the highest air pressures. At 77° C. the difference is appreciable, because at this temperature the vapour pressure is so high that even at the highest air pressures used $p_v/P = 0 \cdot 3$ and the actual expansion suffered by the mixture is uncertain.

From these points, and from a few isolated results at lower temperatures and high air pressures, where the amount of vapour is insignificant, the following table, showing the variation of the supersaturation at the cloud limit with the temperature, has been calculated :—

Table I.

Initial temperature, θ_1.	$\frac{v_2}{v_1}$.	Final temperature. θ_2.	ρ_1 gr./litre.	ρ_2 gr./litre.	$S = \frac{\rho_1}{\rho_2} \cdot \frac{v_1}{v_2}$.
° C.		° C.			
77	1·252	47·0	0·2625	0·0729	2·87
50	1·286	19·1	0·0837	0·0174	3·74
35	1·314	3·2	0·0399	0·00601	5·07
18	1·370	−16·4	0·0154	0·00144	7·80
7	1·375	−26·4	0·00780	0·000636	8·95

In this table the density of saturated water vapour has been found, at temperatures above 0° C. from the steam tables. Below 0° C. it has been deduced from the specific volume, given in c. ft./lb. by Callendar's equation $V - b + c = RT/P$, where $b = 0 \cdot 0160$ and $c = 0 \cdot 4213\,[373 \cdot 1/T]^{10/3}$ at T° A., the saturation pressure in lbs./sq. ins. being determined from the equation

$$\log_{10} P = 21 \cdot 075 - 2903 \cdot 39/T - 4 \cdot 71734 \log T - 0 \cdot 4057\,(c - b)\,P/T,$$

the last term of which is negligible below 0° C.

The values of the saturation pressure so obtained agree very well with the experimental results of Scheele and Heuse, which extend to − 20° C., and they can lead to no serious error in the region − 20° C. to − 40° C.

The results shown in Table I are plotted in fig. 7, A. It will be seen that there is a very great variation in the supersaturation at the cloud limit as the temperature changes. In order to extend the results to higher temperatures, such as are in common use in steam engines and in the steam turbine, use has been made of the fact that the supersaturation must approach unity in the neighbourhood of the critical temperature. This fact and the experimental curve enable us to predict the variation of the supersaturation at intermediate temperatures with considerable confidence, and in fig. 7, B, the interpolation in this region is shown.

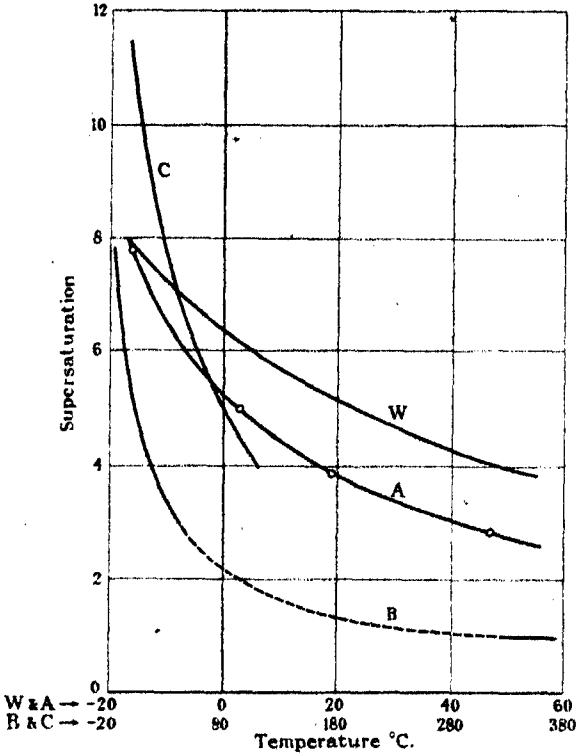


Fig. 7.

§ 9. *The Supersaturation at the Ion Limit.*

The results of experiments, made at different initial temperatures, on the ion limit are shown in fig. 8, where the expansion ratio required to produce condensation on ions is plotted against the initial composition of the expanding mixture p_v/P. From the results the corresponding values of the supersaturation, produced in adiabatic expansion, can be calculated.

Assuming, as in the case of the cloud limit, that the supersaturation at the ion limit is independent of the presence of air, and that it can best be determined by using the results obtained in the experiments in which the air density is so great that the rate of diffusion is very small, the supersaturation at the ion limit can be found. The values given in Table II have been deduced from the points where the extrapolated curves in fig. 8 cut the ordinate $p_v/P = 0$, and

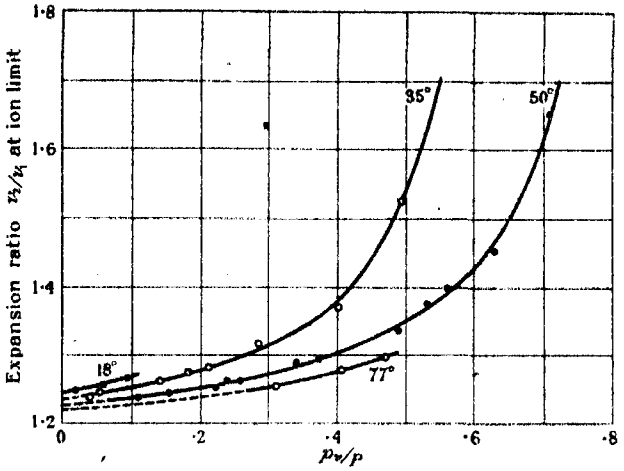


FIG. 8.

Table II.

Initial temperature.	$\frac{v_2}{v_1}$.	Final temperature.	ρ_1 gr./litre.	ρ_2 gr./litre.	$S = \frac{\rho_1}{\rho_2} \cdot \frac{v_1}{v_2}$.
° C.		° C.			
18	1·245	−6·5	0·0154	0·00311	3·98
35	1·235	10·0	0·0399	0·00942	3·44
50	1·226	24·7	0·0837	0·0231	2·96
77	1·218	50·5	0·2625	0·0855	2·52

they are plotted in fig. 9, where the results for the cloud limit are also given for comparison.

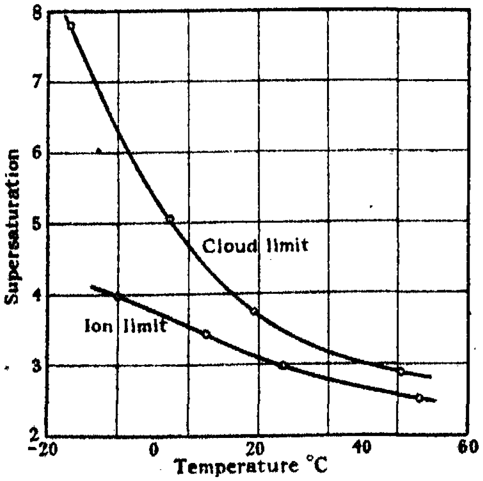


FIG. 9.

In this figure it will be seen that the two curves rapidly approach one another as the temperature rises, and it is evident that there will be serious difficulties in applying the cloud track method at temperatures much above that of the atmosphere. These objections may be seen more directly in fig. 10, in which the expansion ratio required to produce condensation on ions and on associated molecules are shown for different values of the initial temperature of the expanding mixture. Since the expansion ratio at the cloud limit becomes much nearer to that at the ion limit as the temperature rises, the difficulties due to general fogging will be much increased.

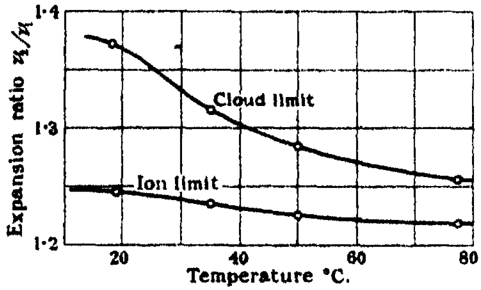


Fig. 10.

It is interesting to apply to these results J. J. Thomson's formulæ

$$\log_e S = \frac{1}{RT}\left(\frac{2\sigma}{r} - \frac{e^2}{8\pi r^4}\right),$$

in which S is the supersaturation of water vapour in equilibrium with a drop of water of radius r, charge e, and surface tension σ, and R, T, are the gas constant per gram and the absolute temperature of the vapour respectively.

For a drop carrying a given charge e, S has its maximum value when

$$\frac{d}{dr}(\log_e S) = 0,$$

i.e., when

$$\frac{2\sigma}{r^2} = \frac{e^2}{2\pi r^5}.$$

This maximum value S_m is therefore given by the formulæ

$$\log_e S_m = \frac{1}{RT}\left(\frac{2\sigma}{r_m} - \frac{\sigma}{2r_m}\right) = \frac{1}{RT}\frac{3T}{2r_m},$$

where $r_m^3 = e^2/4\pi\sigma$, r_m being the particular value of the radius corresponding to S_m.

In order that a drop may grow indefinitely and so become visible, the supersaturation must, initially, be greater than S_m. By calculating r_m from the known value of $e = 4 \cdot 77 \times 10^{-10}$ E.S.U., and the ordinary value for the surface tension, the theoretical value of S_m, the supersaturation necessary just to produce condensation on ions can be found, and in Table III values of this quantity, calculated for different temperatures, are compared with those found by experiment.

Table III.

Temperature.	Theoretical supersaturation.	Experimental supersaturation.
° C.		
−6·5	4·59	3·98
10·0	3·94	3·44
24·7	3·52	2·96
50·5	2·96	2·52

§ 10. *The State of Steam at the Cloud Limit.*

The possibility of the existence of steam in a state of supersaturation in the cylinder of a steam engine was first shown by Callendar and Nicholson,[*] and that it may also exist in the turbine has since been pointed out by other writers. According to Callendar, the steam passes through the turbine so quickly that thermal equilibrium cannot be maintained. It becomes supersaturated, and no appreciable condensation takes place until the cloud limit is reached, when nuclei are produced in enormous numbers. In order to determine the loss in efficiency due to supersaturation, it is necessary to find the state of the steam at the cloud limit, and Callendar[†] has made extensive calculations, assuming as a first approximation that the supersaturation at the cloud limit is eightfold and independent of the temperature. From the results of his calculations the value of the total heat at the cloud limit at different pressures has been plotted in fig. 11, curve W.

That the supersaturation cannot be constant is, of course, evident from the fact that it must be equal to unity at the critical temperature. In order to obtain a better approximation to the truth, therefore, H. M. Martin[‡] applied Kelvin's formula

$$\log_e \frac{p_2}{p_1} = \frac{1}{RT} \cdot \frac{2\sigma}{r}$$

---

[*] Callendar and Nicholson, 'Proc. Inst. Civ. Eng.,' vol. 131, pp. 1–126.

[†] Callendar, 'The Properties of Steam,' p. 246 (1920).

[‡] H. M. Martin, "A New Theory of the Steam Turbine." Reprint from 'Engineering.'

to determine the supersaturation p_2/p_1 at different temperatures. In this formula, p_2 is the pressure of water vapour in equilibrium with a drop of radius r, p_1 the vapour pressure over a flat surface, and σ the surface tension, all at temperature T° A.

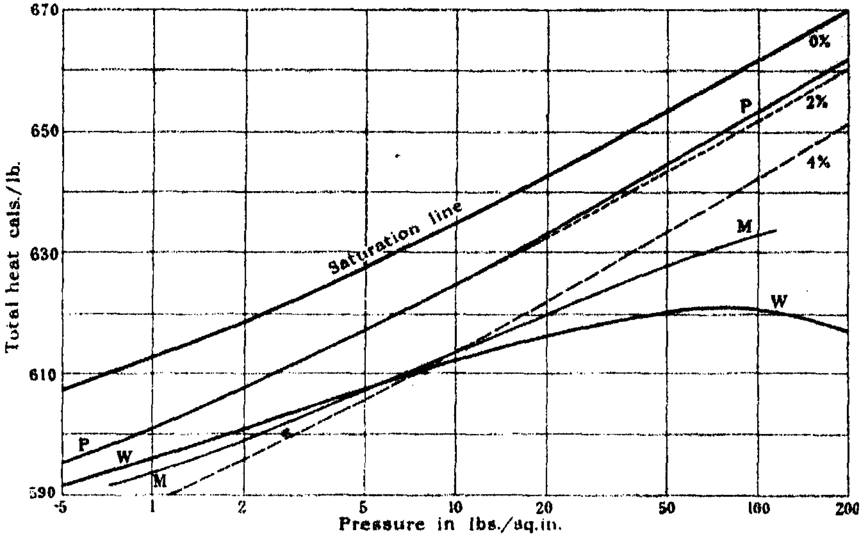


FIG. 11.

Assuming the radius of the nuclei which become effective at the cloud limit to be independent of the temperature, and equal to Callendar's value 5×10^{-8} cm., and taking into account the variation of surface tension with temperature, Martin found a change in the supersaturation from 11·08 at 0° C. to 4·15 at 110° C. His results are shown in fig. 7, curve C.

The value 5×10^{-8} for the radius of the nuclei was, however, calculated on the assumption that the supersaturation at + 20° C. is eightfold, whereas, as has been emphasised, this is the value at − 20° C. Martin's method led to a value of 15 for the supersaturation at − 20° C., and his results were not, therefore, in agreement with the only experimental measurements available. The curve representing the state of the steam at the cloud limit on the total heat-entropy diagram, drawn from his results, is familiar to engineers as the Wilson line. The corresponding curve on the total heat-pressure diagram is shown in fig. 11, curve M.

If we apply Kelvin's formulæ to Wilson's results we obtain the value $6 \cdot 3 \times 10^{-8}$ cm. for the radius of the nuclei at the cloud limit at − 20° C. The results of applying Martin's method using this value of r are shown in fig. 7, curve W, and it will be seen that they are in much better accord with experiment.

The table below shows the state of steam at the cloud limit deduced from the experiments described in this paper. The calculations have been based on Callendar's equation for steam, and, except for a slight modification, necessary since the supersaturation is defined here as a ratio of densities instead of a ratio of pressures, his methods have been adopted. The two definitions lead to values for the supersaturation which are practically indistinguishable at temperatures below 100° C.

From fig. 7 the value of the density ρ_a of steam at the cloud limit at any temperature can be determined, since ρ_a/ρ_s and ρ_s are known. The reciprocal of the density gives the specific volume, and the corresponding pressure is deduced from Callendar's equation for the isothermal

$$P(V - b + c) = \text{constant}.$$

When the pressure is measured in lbs./sq. ins. and the volume in c. ft./lb., $b = 0 \cdot 0160$, and c at any given temperature can be found from the steam tables.

The pressure and temperature being known, the total heat in calories/lb. can be determined from the equation

$$H = S_0T - SCP + 464, \text{ where } S_0 = 0 \cdot 47719.$$

SC is given in the steam tables for temperatures above 100° C., and can be deduced for lower values from the equation

$$SC = \frac{144}{1400}\left[\frac{13}{3}c - b\right].$$

Similarly the entropy in lb. calories/degree is given by

$$\phi = 1 \cdot 09876 \log_{10} \frac{T}{373 \cdot 1} - 0 \cdot 25356 \log_{10} \frac{P}{14 \cdot 689} + 1 \cdot 763 - \frac{ancP}{T}.$$

In the table the suffix "*ss*" indicates that the quantity to which it is applied is measured at the cloud limit. Similarly the suffix "*s*" refers to measurements at saturation.

Table IV.—State of Water Vapour at the Cloud Limit.

Temperature.	P_s.	Supersaturation = S.	V_s.	V_{ss}.	P_{ss}.	H_{ss}.	ϕ_{ss}.
° C.							
0	0·0892	5·40	3275·9	606·5	0·4811	594·02	1·9899
10	0·1789	4·40	1693·8	384·9	0·7855	598·72	1·9529
20	0·3399	3·70	922·19	249·2	1·254	603·32	1·9174
30	0·6162	3·25	525·8	161·8	1·992	607·85	1·8819
40	1·070	3·02	312·4	103·5	3·214	612·30	1·8441
50	1·789	2·81	192·7	68·59	4·995	616·67	1·8098
60	2·887	2·53	122·9	48·59	7·250	620·98	1·7821
75	5·586	2·22	66·20	29·82	12·31	627·21	1·7448
90	10·161	1·99	37·81	19·00	19·99	633·19	1·7074
100	14·69	1·86	26·79	14·40	26·96	637·02	1·6872
120	28·81	1·66	14·27	8·595	47·11	644·22	1·6450
140	52·48	1·51	8·143	5·392	78·91	650·70	1·6068
160	89·80	1·39	4·923	3·542	122·7	656·85	1·5755
180	145·6	1·29	3·127	2·408	185·6	662·24	1·5423
200	225·2	1·20	2·074	1·728	266·0	667·58	1·5205

From these results curve P, fig. 11, has been drawn. The dotted lines in the same diagram are lines of constant wetness, and it will be seen that curve P coincides with a line of 2 per cent. wetness almost exactly at pressures below 20 lbs./sq. ins., the region of most importance in turbine theory. This enables the results to be applied with a great simplification in the calculations.

It is worthy of note that from his analysis of the pressure distribution in the turbines of s.s. "Mauretania," Callendar* concluded that the line W in fig. 11 was too low at high pressures, and in its place he proposed that the line of 3 per cent. wetness should be taken as being best in accord with the experimental facts. When the nature of the experimental data at his disposal is considered, the agreement with the results of the present experiments is satisfactory.

In general we may conclude that the losses of efficiency in nozzles and turbines, due to supersaturation, have been over-estimated. The results just given enable an upper limit to these losses to be calculated for any particular case, and, as will be shown in another place, for rapidly expanding steam this limit agrees with the values deduced from experiment.

Summary.

(1) An apparatus is described whereby the supersaturation required to produce condensation on ions and on associated molecules in the presence of

---

* Callendar, 'Properties of Steam,' p. 249.

dust-free air, over a range of temperature from − 25° C. to 50° C., has been measured.

(2) The effect of the air on the phenomena has been investigated, and it has been shown to be of importance not only by determining the amount of water vapour condensed per unit volume for a given degree of supersaturation, but also by preventing the evaporation of water from the walls of the expansion chamber.

(3) By an extrapolation the supersaturation necessary to produce cloudy condensation at temperatures up to the critical point has been estimated, and the state of steam at the cloud limit at different temperatures calculated. The importance of the results in connection with the thermodynamic theory of the steam turbine is explained.

(4) It is shown that, apart from the experimental difficulties to be met in working at higher temperatures, atmospheric temperature gives the best conditions for the cloud method of investigating atomic phenomena.

It is with great pleasure that I acknowledge my indebtedness to Prof. C. T. R. Wilson, who first suggested the problem, and who has given me the benefit of his advice and encouragement throughout the course of this work. My thanks are due to the Department of Scientific and Industrial Research for a grant.

The Effect of Temperature on the Viscosity of Neon.

By R. S. Edwards, B.Sc., A.R.C.S., D.I.C.

(Communicated by H. L. Callendar, F.R.S.—Received March 27, 1928.)

1. *Introduction.*

The experiments described in this paper have been carried out to obtain accurate data with respect to the variation of the viscosity of neon with temperature, over a range of several hundred degrees Centigrade. This has been rendered more desirable in view of the recent publication of a new theoretical formula for the variation of viscosity of a gas with temperature by J. E. Lennard-Jones,* which he has applied with conspicuous success to various gases. In the case of neon, however, it was not possible to make a test of the formula as the experimental data are so sparse, only one observer, A. O. Rankine,† having made measurements upon its viscosity, and then at only two temperatures. Accordingly, as neon should provide a crucial test between Lennard-Jones' formula and that of Sutherland, the method already applied by the author‡ to the case of air was suitably modified for use with a gas of which the quantity available was strictly limited and values of the viscosity obtained from 444·5° C. downwards. Owing, however, to the increase in the quantity of neon required as the temperature was reduced, it was not found possible to take observations at a temperature lower than — 78·4° C. Unfortunately, it is only at low temperatures that the difference between the two formulæ becomes marked, and hence the results obtained have not been sufficient to provide a really crucial test between them. The determination of the viscosity over a large range does, however, for the first time, provide sufficient data for a reliable comparison to be made of the molecular properties of neon, as determined from its viscosity with the values of the same properties deduced from other sources.

2. *Theory of the Method.*

The principle of the method was identical with that used in the experiments on air. Briefly, the neon was enclosed in a fairly large bulb and allowed to escape through a capillary tube against a constant pressure, both the bulb and capillary being surrounded by suitable jackets for maintaining the various steady temperatures at which observations were taken. Hence the formula to calculate the viscosity is the same as that given in the previous paper (*loc. cit.*). This was

---

* 'Roy. Soc. Proc.,' A, vol. 106, p. 441 (1924) and subsequent papers.

† 'Roy. Soc. Proc.,' A, vol. 83, p. 516, and vol. 84, p. 181 (1910).

‡ 'Roy. Soc. Proc.,' A, vol. 117, p. 245 (1927).

$$\eta_t = \frac{\pi r^4 \tau}{8Vl} \frac{p_0(1 + 4\xi_t/r)}{\log_e \{(P_2 + 2p_0)P_1/(P_1 + 2p_0)P_2\}}, \tag{1}$$

where

t = degrees Centigrade, T = corresponding absolute temperature.
η_t = viscosity at t degrees Centigrade.
r = radius.
l = length capillary.
V = volume of bulb.
τ = time of flow.
p_0 = steady external pressure outside bulb.
P_1 = initial excess pressure in bulb over the external pressure.
P_2 = final excess pressure in bulb over the external pressure.
ξ_t = coefficient of slip at t° C.

The term $\pi r^4/8Vl$ is independent of the temperature if the bulb and tube are of the same material; hence, writing K for this constant term, we have

$$\frac{\log_e \{(P_2 + 2p_0)P_1/(P_1 + 2p_0)P_2\}}{\tau p_0} = \frac{K(1 + 4\xi_t/r)}{\eta_t}. \tag{2}$$

Therefore, if measurements are made of the four quantities P_1, P_2, p_0, and τ at two different temperatures t_1 and t_2, then provided the radius of the capillary is known, we have at once the ratio η_{t_2}/η_{t_1}, which gives, for different values of t_2, the variation of the viscosity with temperature.

3. *Description of Apparatus.*

In designing the apparatus, apart from the necessity of keeping the total amount of neon as small as possible, two conditions had to be complied with. These were, to arrange for the neon which escaped from the bulb during a run to be returned in order that the series of operations could be repeated, and, secondly, to maintain a steady pressure for the closed volume into which the neon escaped. Provision had also to be made for the preliminary exhaustion and filling with the pure gas. Fig. 1, which is a diagrammatic sketch of the apparatus, shows the arrangement adopted.

The requisite steady pressure beyond the tap T_3 was attained by the use of the aspirator P, which was made in the shape shown in order to have a large cross section and comparatively small volume. The cross section of the liquid surface in P (water was actually used) was such that the depression in the head due to the neon passing into P was negligible during an observation. The neon was returned to the bulb B after a run by way of the manometer M, the mercury in the arms *a* and *b* being lowered beneath their junction for this purpose.

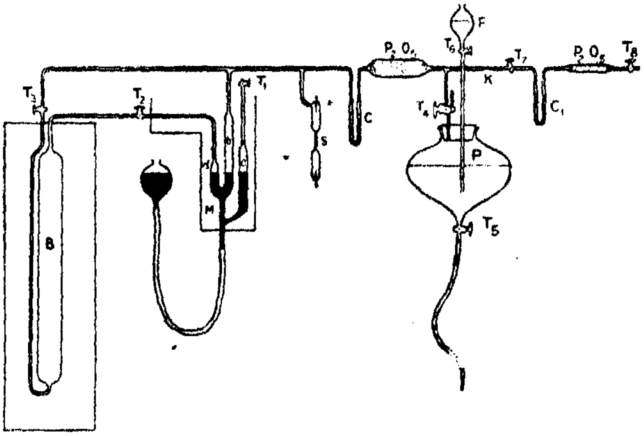


Fig. 1.

A large P_2O_5 tube was inserted between P and the bulb in order to prevent water vapour reaching the bulb. The charcoal tube C was kept continually surrounded by liquid air throughout the whole series of measurements. The small discharge tube S was used as a qualitative test of the purity of the neon. This test was fairly sensitive as it was found that a small 4-volt induction coil was sufficient to obtain a discharge in the pure gas, but that the addition of 1 to 2 per cent. of air was sufficient to stop the discharge. The dropping funnel F was used to return to the aspirator the water run off during an experiment, a procedure which simultaneously restored the pressure in the apparatus to its initial value.

The bulb B and the capillary tube were the same as used in the previous work (*loc. cit.*), as were also the water jacket around the manometer, the arrangements for reading the mercury levels and for their fine adjustment, and consequently need no further description. The following list gives the dimensions of the apparatus :—

Volume of bulb = 606·3 c.c. at 15° C.
Volume of portion of bulb not enclosed by jacket = 0·80 c.c.
Volume of tube from end capillary to tap T_3 = 0·90 c.c.
Radius capillary = 0·013 cm.
Length capillary = 76·7 cms.
Volume from tap T_2 to fixed level in limb *a* = 5·10 c.c.

Readings were taken at six temperatures, viz., − 78·4°, 15°, 100°, 184·4°, 302°, 444·5° C. The jackets employed for the determinations at room and the higher temperatures were the same as previously described (*loc. cit.*), and the same methods for obtaining the temperatures were also adopted. At − 78·4° C. a large Dewar flask filled with a mixture of solid CO_2 and alcohol was used as the jacket. To enable the state of the mixture to be seen a strip of glass had been left unsilvered throughout the length of the vessel. The temperature was determined by a platinum thermometer in the usual manner. Determinations of the viscosity at liquid air temperature would have been most desirable, but the amount of extra neon required to sustain atmospheric pressure in the apparatus due to the large drop in temperature rendered this impossible.

4. *Purification of the Neon.*

Before describing the method of operation of the apparatus it will be necessary to describe the purification of the neon. This has been regarded as a matter of considerable difficulty because of the trouble caused by helium as an impurity. The method adopted was a simple modification of that introduced by Sir W. Ramsay* and elaborated by H. E. Watson.† This process, one of fractional distillation over charcoal surrounded by liquid air, enabled Watson with the aid of an elaborate system of density measurements to separate completely large volumes of mixed neon and helium. In the present case, however, the neon obtained from the British Oxygen Company was much purer, 98 per cent. being the figure claimed by the Company. Consequently, the procedure was adopted of using only the middle portion of the gas coming over during the fractional distillation, rejecting the initial and final portions. This greatly simplified the apparatus, a sketch of which is given in fig. 2. Care was taken in

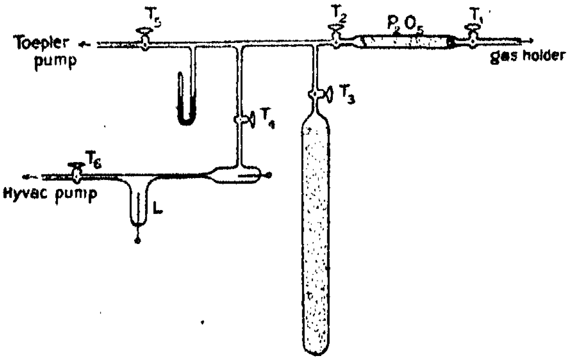


FIG. 2.

---

* 'Roy. Soc. Proc.,' A, vol. 76, p. 111 (1905).

† 'J. Chem. Soc.,' vol. 97, p. 810 (1910).

assembling to avoid "dead spaces" and to make the connecting tubes as short as possible in order to avoid loss of neon.

The operations were carried through as follows: With all the taps except T_1 open, the system was exhausted to the limit of a Hyvac pump, the charcoal tube being heated thoroughly to remove absorbed air. T_4, T_5, T_6 were then closed and the liquid air placed round the charcoal bulb, which had a capacity of about 750 c.c. of neon under atmospheric pressure. The neon was then brought into contact with the charcoal, a few cubic centimetres at a time, by opening and closing T_1 and T_2 alternately until the charcoal was saturated. By this process the charcoal and gas holder were separately brought into contact with the P_2O_5 tube, thus preventing a rush of gas from the holder, which, had it occurred, would have made the necessary filling of that vessel with water very difficult. The P_2O_5 also served to prevent any water vapour reaching the charcoal.

Since helium is practically unabsorbed by charcoal surrounded by liquid air, almost the whole of the helium content, estimated to be not more than 25 to 30 c.c. in 750 c.c. of neon, was contained in the unabsorbed gas above the charcoal. On the other hand, any air and water vapour were held by the charcoal and exerted no appreciable pressure. Thus the pressure (ultimately made atmospheric) above the charcoal was due to the helium impurity and unabsorbed neon. T_2 was then closed, T_5 opened, and about 100 c.c. (measured at N.T.P.) of gas were pumped off by the Toepler and collected in a mercury aspirator connected to the delivery end of the pump. To do this, T_3 was opened until a convenient pressure as measured by the vacuum manometer was attained, then closed and the gas pumped off. T_3 was then permanently opened and the bulk of the neon pumped off and collected in a separate aspirator. Pumping was stopped when the pressure had fallen to 1 cm.; this ensured that no air or water vapour came over with this portion, which was taken as pure. The amount of pure neon obtained was about 500 c.c., the original quantity being about 750 c.c. The process was then repeated until sufficient of the pure neon, about 1200 c.c., had been obtained, again expressed in terms of N.T.P.

During the course of the experiments, as a test of the purity of the neon, a portion was redistilled in the same manner and mixed with an equal volume of once distilled neon; measurements of the viscosity were then taken and compared with those from once distilled neon. It was found from these observations (discussed more fully later) that a single distillation was sufficient to purify the neon, and therefore additional supplies of neon, occasioned from time to time by air leaks and other causes, were distilled once only. The test just described was made with a 50 per cent. mixture and not with an entirely twice distilled supply in order to avoid wastage of neon as far as possible.

An attempt was made to examine spectroscopically the portion of the gas containing the helium. For this purpose the tap T_4 was opened and a portion of the unabsorbed gas was introduced into the discharge tube. This was then pumped down to a pressure of about 0·1 mm. and the cathode glow round the electrode L examined. The test was made in this manner as it provided the most favourable conditions for observing helium in the presence of neon. The attempt was made to identify the helium line λ 5875 A.U., which is close to the neon line λ 5882 A.U., and which should cause the latter line to appear as a doublet. According to Watson (*loc. cit.*), at least 10 per cent. of helium is necessary before the doublet appears, and this was confirmed in the present case, as it was found necessary to redistil the initial portion of the distillate before the doublet could be observed. This test, therefore, though of no use as a guide to the purity of the neon, served to show that the method concentrates the helium in the initial portion of the distillate.

5. *Experimental Method.*

Before filling with neon, the apparatus shown in fig. 1 was thoroughly exhausted and wherever possible the parts were well heated during the process. Meanwhile, the aspirator had been filled with recently boiled distilled water; to do this T_4 was turned to connect P with the atmosphere and the vessel then filled by displacement.

C_1 was then surrounded by liquid air, the gas holder attached to T_8, and the neon passed into the apparatus in small volumes at a time by opening and closing T_7 and T_8 alternately. The P_2O_5 and C_1 removed the water vapour and any air that may have been admitted into the neon during the transference operations. When the pressure in the apparatus had reached atmospheric, T_4 was turned to connect P to the rest of the apparatus, and with T_5 open neon was drawn into the aspirator until the amount was adjudged sufficient. T_5 was then closed and finally the constriction at K was sealed off.

It was found, however, that air was gradually given off from the water although precautions had been taken to prevent this, and finally it became necessary to keep the charcoal tube C continuously surrounded by liquid air. This prevented the air reaching the bulb, which was all that was required, its presence in the aspirator being of no importance.

The procedure in taking observations was as follows:—The mercury was raised above the junction of the limbs *a* and *b* of the manometer, T_3 closed,

T_4 turned to connect the aspirator to the rest of the apparatus and T_5 opened. This reduced the pressure on the exit side of the capillary tube to a value below atmospheric dependent on the head of water in P. The level in a was then adjusted to a fixed position, and when the pressure in the bulb was steady, as noted by observation of the level in a, T_2 was closed, T_3 opened and the stop-watch started simultaneously. During the run the mercury levels in a, b, c, were noted with a cathetometer microscope reading to 0·1 mm. T_1 was open to the atmosphere throughout. At the end of the run, T_3 was shut and the watch stopped simultaneously. T_5 was quickly closed and the mercury lowered until the levels in b and c were approximately in the position they were expected to occupy when T_2 was opened. T_2 having been opened, the level in a was adjusted to the position it occupied at the start and the level in c noted. Water was then run into P until the levels of the limbs a and b were the same, and the mercury lowered beyond their junction. Finally, water was run into the aspirator until the pressure in the whole apparatus was brought up to its original value. By either varying the head of water, or by altering the initial pressure throughout the apparatus (by running more or less water into the aspirator), it was possible to change the initial difference of pressure between wide limits.

6. *Calculation of Results.*

The pressure differences, measured as described in the previous section, need some correction before substitution in formula (2). P_1 is measured directly by the observed difference of level between the limbs a and b, noted during the run. The value of p_0 is also given directly, being equal to atmospheric pressure less the difference of level of the limbs b and c, again noted during the run. The value of P_2 is calculated in the following manner: when the tap T_2 is opened, a volume V of gas at a pressure $(P_2 + p_0)$ and temperature T is mixed with a volume v at a pressure $(P_1 + p_0)$ and temperature T′, where v = volume from tap T_3 to the fixed level in limb a, and T′ is the temperature of the manometer. Hence, by the gas laws, we have

$$V(P_2 + p_0)/T + v(P_1 + p_0)/T' = (p_0 + P)(V/T + v/T'),$$

i.e.,

$$P_2 = P - vT(P_1 - P)/VT'$$

or

$$P_2 = P_1 - (1 + vT/VT')(P_1 - P),$$

where $(P + p_0)$ is the uniform pressure attained on mixing. The quantity $(P_1 - P)$ is given directly by the difference of level in the limb c at the commencement and end of a run, hence P_2 is known. A small correction to this value of P_2, necessitated by the presence of the small volume between the end of the capillary and the tap T_3, has already been discussed in the previous paper (*loc. cit.*), and hence it will be sufficient to give the table of values.

Excess pressure in bulb.	Correction.
cms.	cms.
2	0·0030
3	0·0045
4	0·0060
5	0·0074

Formula (2) expressed in practical units is

$$\frac{\log_e \{(P_2+2p_0)\,P_1/(P_1+2p_0)\,P_2\}}{\tau p_0} = \frac{15\pi r^4\, g\rho_0\,(1+4\xi_t/r)}{4\cdot 606 V l \eta_t},$$

where the pressures are measured in centimetres of mercury at 0° C. and the time in minutes. ρ_0 = density of mercury at 0° C.

Therefore we have

$$K = 15\pi r^4\, g\rho_0/4\cdot 606 V l.$$

P_1 and P_2 were measured at t', the temperature of the manometer, and were not reduced to 0° C., the slight error introduced, about 1/15000, in the logarithmic term being negligible. The expression for the coefficient of slip $(1+4\xi_t/r)$ was treated in the same way as previously described (*loc. cit.*).

7. *Results.*

No attempt has been made to obtain absolute values for the viscosity of neon using the dimensions of the apparatus, but the alternative plan has been adopted of calibrating the apparatus in terms of a standard value for the viscosity of air. The standard chosen was that given by Millikan,* viz.:—

$$\eta_t = 0\cdot 0001824 - 0\cdot 000000493\,(23 - t) \text{ c.g.s. units,}$$

over the range 23° to 12° C., which gives

$$\eta_{15} = 0\cdot 00017846 \text{ c.g.s. units.}$$

This method has the advantage that a uniform basis of comparison with the results of other observers is attained, and should any alteration of the standard be found necessary in the future, all other results can at once be readjusted.

Table I gives the list of values of $K\,(1+4\xi_t/r)/\eta_t$ obtained for the purpose of calibrating the apparatus in terms of Millikan's air value. The air was dried and freed from CO_2 in the usual way. As a table showing a typical series of observations was included in the previous paper (*loc. cit.*), only the results are quoted here.

---

* Millikan, 'Ann. d. Physik,' vol. 41, p. 759 (1913).

Table I.

Temperature	$\frac{K(1+4\xi_t/r)}{\eta_t} \times 10^4$	$\frac{K(1+4\xi_{15}/r)}{\eta_{15}} \times 10^4$	Temperature	$\frac{K(1+4\xi_t/r)}{\eta_t} \times 10^4$	$\frac{K(1+4\xi_{15}/r)}{\eta_{15}} \times 10^4$
° C.			° C.		
16·86	4·798	4·823	18·36	4·781	4·826
17·58	4·759	4·796	18·36	4·772	4·817
17·58	4·758	4·795	18·32	4·787	4·832
17·58	4·800	4·835	16·97	4·777	4·804
18·24	4·771	4·815	17·05	4·792	4·820
18·25	4·771	4·815	17·48	4·772	4·806
18·26	4·766	4·810	17·48	4·768	4·802
18·27	4·767	4·811	17·48	4·809	4·843
18·53	4·790	4·837	17·48	4·761	4·795
18·50	4·790	4·837	17·54	4·813	4·847

The third column gives the values of $K(1 + 4\xi_t/r)/\eta_t$ corrected to 15° C. by the use of Millikan's formula. The mean value thus determined is $K(1 + 4\xi_{15}/r)/\eta_{15} = 4{\cdot}8183 \times 10^{-4}$ c.g.s. units. The value of $(1 + 4\xi/r)$ at 15° C. is 1·0025; therefore $K = 8{\cdot}5789 \times 10^{-8}$, and since the essential dimensions of the apparatus were unaltered from those of the previous work (*loc. cit.*), the two values for K should agree. The value previously obtained was $K = 8{\cdot}5819 \times 10^{-8}$, which agrees with the above to approximately one part in three thousand. This furnishes a satisfactory proof of the applicability of the modified form of apparatus.

Table II gives the observations on neon. The readings taken, about 15° and 100° C., were corrected to these two temperatures using Rankine's value (*loc. cit.*) for the rate of change of viscosity with temperature over this range. These having been determined, a provisional value of Sutherland's constant was calculated, and at the other temperatures the observations were reduced by the use of his formula. This procedure was sufficiently accurate, as with the exception of the readings at room temperatures the variation was almost negligible.

Table II.

Temperature (° C.).	−78·4	15·0	100·0	184·2	302·0	444·5
$K\,(1 + 4\,\xi_t/r)/\eta_t \times 10^4$ c.g.s. units.	3·643	2·804	2·360	2·073	1·768	1·570
	3·641	2·788	2·356	2·078	1·767	1·547
	3·633	2·799	2·373	2·076	1·772	1·551
	3·614	2·812	2·350	2·080	1·772	1·569
	3·617	2·799	2·363	2·057	1·781	1·570
	3·638	2·793	2·352	2·076	1·761	1·541
	3·674	2·802	2·357	2·070	1·781	1·540
	3·624	2·816	2·371	2·074	1·775	1·563
	3·620	2·792	2·369	2·063	1·774	1·555
		2·819	2·356	2·068	1·762	1·548
		2·788	2·353	2·060	1·772	
		2·782	2·362	2·062		
		2·812		2·069		
		2·806		2·080		
				2·052		
Means	3·634	2·801	2·360	2·069	1·771	1·555

The values of the factor $(1 + 4\xi_t/r)$ were,

Temperature (° C.).	−78·4	15·0	100·0	184·4	302·0	444·5
$(1 + 4\,\xi_t/r)$	1·0027	1·0043	1·0059	1·0076	1·0099	1·0125

Substituting these factors, the following mean values were obtained for the viscosity :—

Table III.

Temperature (° C.).	−78·4	15·0	100·0	184·4	302·0	444·5
$\eta_t \times 10^4$ c.g.s. units	2·367	3·076	3·656	4·177	4·901	5·584

Table IV gives a comparison of observations taken with (*a*) the commercial neon, (*b*) once distilled neon, and (*c*) 50 per cent. redistilled + 50 per cent. once distilled neon (*cf.* § 4), taken at 15° C.

Table IV.

Neon used.	$K\,(1 + 4\xi/r)/\eta_{15} \times 10^4$ (mean).	Number of observations.
Commercial neon	2·913	7
Once distilled	2·802	10
50 per cent. redistilled + 50 per cent. once distilled	2·797	4

8. *Accuracy of the Results.*

The degree of accuracy to be attained by this method of determining the viscosity of a gas was fully discussed in the previous publication (*loc. cit.*). It was there shown that the expression for the percentage accuracy was

$$d\eta/\eta \propto dx/x \log_e x,$$

where x denotes P_1/P_2. It follows, therefore, that the greater x the greater the accuracy attained, provided P_2 is not at the same time made so small that the error in measurement is increased to an extent outweighing the advantage gained by the increase in x.

The average value of P_2 was about 3·0 cms., and as the cathometer read to 0·1 mm., the error in measurement can be estimated to be of the order of 0·5 per cent. The high viscosity of neon prevented, however, x being made much greater than 2·0; had this been done, the time of flow would have been unduly long. Hence, the accuracy of individual readings may be taken as of the order of 0·8 per cent. On this basis, the accuracy of the means of the groups of results would be approximately one part in five hundred.

9. *Discussion of Results.*

As already stated, very little comparison with previous results can be made owing to the scarcity of data. Table V gives the comparison between the results of A. O. Rankine (*loc. cit.*), reduced to 15° C. and 100° C., and expressed in terms of Millikan's value for air instead of that used by Rankine, and the corresponding present results.

Table V.

$\eta_t \times 10^4$ c.g.s. units.	Rankine.	Edwards.	Difference.
			per cent.
η_{15}	3·069	3·076	0·23
η_{100}	3·622	3·656	0·93

Before a comparison can be drawn between the two sets of results, an estimation of the purity of the neon used in the present experiments must be made. That used by Prof. Rankine had been prepared by H. E. Watson (*loc. cit.*) for his determination of the atomic weight of neon, and hence was of undoubted purity. While no direct test of purity by density determinations as performed by Watson has been carried out in the present case, comparative measurements of the viscosity have been utilised to serve the same purpose. The figures in

Table IV show that the value of the quantity K $(1 + 4\xi/r)\eta_{15}$ was unchanged within the limits of experimental error for once distilled, and for 50 per cent. redistilled + 50 per cent. once distilled neon. This provides evidence that the process described in § 4 sufficiently removes the helium impurity in a single distillation.

Regarding, therefore, the gas as pure, the agreement between the present result and that of Prof. Rankine at 15° C. is very satisfactory. In a later paper,* Prof. Rankine has estimated a correction to his value for η_{100}, which would make the result $3 \cdot 663 \times 10^{-4}$ c.g.s. units, which agrees well with the new determination. This correction involved a change in Sutherland's constant from 56, as originally given by Rankine, to 69.

Sutherland's formula for the viscosity of a gas, $\eta_T = kT^{3/2}/(S + T)$, can be tested very easily by plotting $T^{3/2}/\eta_T$ against T, which should give a straight line, and S (Sutherland's constant) is given by the negatived intercept on the T axis. This has been done as shown in fig. 3, from which it will be seen that, with the exception of the determination at $-78 \cdot 4°$ C., the values all fit the mean straight line closely, the maximum divergence being $0 \cdot 7$ per cent. The value of S thus

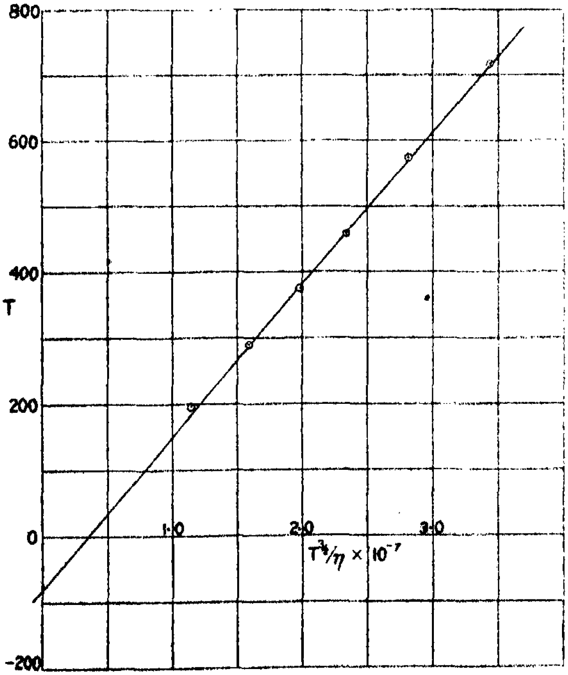


FIG. 3.

---

* Rankine and Smith, 'Phil. Mag.,' vol. 42, p. 601 (1921).

determined is 80, which agrees as well as can be expected with Prof. Rankine's value of 69 when allowance is made for the difference in the range of temperatures covered. The value of the viscosity at $-78 \cdot 4°$ C. as determined by Sutherland's formula is definitely low, a result in agreement with the observations of a number of workers on various permanent gases at low temperatures.

The formula of Lennard-Jones is $\eta_T = kT^{3/2}/(T^x + S)$, where S = constant independent of the temperature and analogous to that of Sutherland, and $x = (n-3)/(n-1)$, where n = the index of the repulsive force between the molecules. For n equal to infinity, the formula reduces to that of Sutherland. Unfortunately Lennard-Jones' formula cannot be tested by any simple graphical method; therefore, the somewhat laborious procedure adopted by Lennard-Jones has been followed, but the results do not show any decided superiority for either formula over the greater part of the range of temperatures investigated. At the lower temperatures, however, the combination $S = 2 \cdot 4$, $n = 14\frac{1}{2}$ was definitely superior. This indicated that had it been possible to obtain values of the viscosity at still lower temperatures, a definite result in favour of Lennard-Jones' formula would have been obtained, a view entirely in accordance with his calculations from the results of Schmidt* for argon. It appears, therefore, that Sutherland's formula is quite satisfactory for temperatures well above the liquefying point, but as that point is approached Lennard-Jones' becomes superior. It is interesting to note that a value of about 14 for n has been found the most suitable for both helium and argon by Lennard-Jones (*loc. cit.*), a result to be expected from the similarity in structure of these gases.

In conclusion, the author wishes to express his indebtedness to Prof. A. O. Rankine for the constant interest in, and frequent advice he has given throughout the course of, this work.

Summary.

(1) The ratios of the viscosities of neon at various temperatures between $-78 \cdot 4°$ and $444 \cdot 5°$ C. have been determined, using a constant volume method and suitable jackets to obtain definite temperatures.

(2) The actual values of the viscosity in terms of a standard air value have also been determined, and their representation by the theoretical formulæ of Lennard-Jones and Sutherland has been discussed.

---

* Schmidt, 'Ann d. Physik,' vol. 30, p. 309 (1909).

The Homogeneous Reaction between Hydrogen and Oxygen.

By C. H. Gibson and C. N. Hinshelwood.

(Communicated by Sir Harold Hartley, F.R.S.—Received March 31, 1928.)

When a mixture of hydrogen and oxygen is heated in a silica vessel at a temperature below about 520° C., the slow reaction which occurs takes place almost entirely on the walls of the vessel, is not very much influenced by pressure, and has a small temperature coefficient. In the next 50° above this, a reaction in the gas phase comes into prominence, the rate of which depends on the concentration of hydrogen and oxygen in such a way as to indicate a reaction of very "high order." The temperature coefficient is high and increases both with temperature and pressure, and the gas-reaction, unlike the surface reaction, is "autocatalysed" by steam. The gas reaction is retarded by an increase in the surface exposed to the gas. The retardation must be due either to the destruction by a surface decomposition of something which catalyses the principal reaction, or to the interruption of reaction chains, in a manner which has already been discussed.*

The homogeneous reaction between hydrogen and oxygen is interesting if only because it is one of the most fundamental reactions from the purely chemical point of view, but it appears also to be specially worthy of further investigation because it is kinetically of so remarkable a nature.

A large number of experiments on the kinetics of the reaction have now been made, porcelain being substituted for silica vessels. The desirability of using vessels of some material other than silica is obvious, for the results of the former investigation might have been determined, partly at least, by peculiar catalytic qualities of the silica, since, although the reaction takes place principally in the gas, the surface plays some part in the mechanism. The new results, however, show that in general the reaction in porcelain vessels is of exactly the same nature as that which occurs in silica vessels. Experiments conducted in porcelain vessels have the advantage of being more regular and comparable one with another; thus the finer points about the reaction appear more clearly.

The influence of various inert gases on the reaction has been investigated, and the results throw some interesting light on the theory of reaction chains, and on the conditions affecting the transfer of internal energy from one molecule to another in collisions.

---

* 'Roy. Soc. Proc.,' A, vol. 118, p. 170 (1928).

Apparatus.—The apparatus was similar to that previously described, except that specially made bulbs, of about 200 c.c. capacity, of Worcester porcelain were used instead of silica bulbs. They were admirably gas-tight, and except that they were a little fragile, were well adapted in every way for high temperature work with gases. They were joined to the glass part of the apparatus by flanges.

"Order" of Reaction: Influence of Surface.

The low order of the reaction at 506° C. is shown by the following experiments :—

Initial pressure of H_2.	Initial pressure of O_2.	Time for 5 per cent. reaction.
mm.	mm.	
(1) 200	100	61′ average 67′
(3) 200	100	72′
(2) 400	200	44′

These give 1·6 for the order.

At 529° C. the order rises to over 2 and above this point increases very rapidly.

Table I contains a summary of experimental results for equivalent quantities of hydrogen and oxygen at different temperatures. The same bulb was used in all these experiments, but in some of them the surface was greatly increased by filling the bulb with fragments of broken porcelain supplied for the purpose by the manufacturers. The time required for half the mixture to react (t_{50}) is recorded as the best empirical measure of the rate of reaction.

The following facts are evident from the table :—

(*a*) At 549° C. and the higher temperatures the influence of pressure on the rate of reaction is remarkably great with the unpacked bulb.

The reaction has not a constant order: if log p is plotted against log t_{50}, the line obtained is not straight but curves in such a way as to indicate that the apparent order increases with pressure. At 559°, for example, the slope of the curve indicates an order changing steadily from 3·15 at 200 mm. H_2 to 5·65 at 400 mm.; and at 569° the order changes from 3·4 in the neighbourhood of 150 mm. to 6·4 in the neighbourhood of 400 mm.

(*b*) In the packed bulb the characteristic behaviour of the surface reaction is to be observed at all temperatures except the highest: namely, that t_{50} does not vary with pressure—except for the small irregular variations which are inevitable in dealing with surface reactions.

Table I.—Time of Half Change for Equivalent Mixtures of Hydrogen and Oxygen at Various Temperatures.

	Temperature.	Pressure of hydrogen.								
		400.	350.	300.	275.	250.	225.	200.	150.	125.
	° C.	′ ″	′ ″	′ ″	′ ″	′ ″	′ ″	′ ′	′ ″	′ ″
Unpacked bulb	506	440 0								
Packed bulb		60 0								
Unpacked bulb	529	126 0						460 0		
Packed bulb		31 0						29 45		
Unpacked bulb	549	38 0		75 40				250 0		
Packed bulb		21 10		23 30				19 5		
Unpacked bulb	559	8 15	15 25	28 5		50 30		73 30	160 0	
Packed bulb		12 47	14 48	14 0		15 40		11 25	11 30	11 30
Unpacked bulb	569	2 7	4 42	7 18	10 40	22 30		39 45	84 20	
Packed bulb		9 6		10 30				10 30	10 20	
Unpacked bulb	576	*	*	2 38		6 25	11 2	17 0		
Packed bulb			7 16	7 10		7 45		9 20	9 15	
Unpacked bulb	582	*	*	*	*	*	*	*	*	
Packed bulb		3 58	5 22	5 43		5 14		5 33	6 54	
Unpacked bulb	590	*	*	*		*		*		
Packed bulb		*	2 24	3 0		4 4		4 18		

* Explosion.

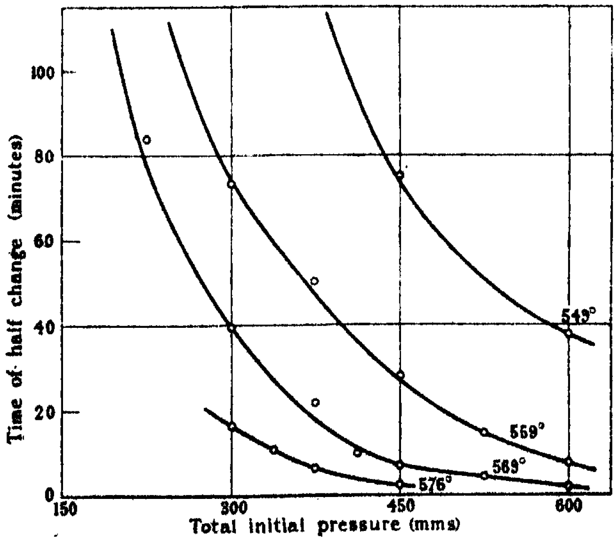


Fig. 1.—Influence of total initial pressure on time of half reaction for equivalent proportions of oxygen and hydrogen in the unpacked bulb.

At low temperature the total rate of reaction is greatly increased by the extra surface. At high temperatures and pressures the rate does not merely become independent of the surface as would be expected when the gas reaction tended to predominate, but is actually much smaller in the packed bulb, showing that the surface retards the gas reaction at the same time as it offers greater facility for the heterogeneous reaction.

The actual speeds measured in the porcelain vessel are smaller than in the silica vessels previously used. The difference is not very great; equal speeds are attained at temperatures differing by about 10° for the two kinds of vessel.

Temperature Coefficient.

When log t_{50} is plotted against the reciprocal of the absolute temperature in the usual way, the line obtained has a variable slope, owing to the transition from surface reaction to gas reaction. Some results are shown in fig. 2. In the neighbourhood of 830° absolute the increase of rate for 10° is 4·3 when the initial pressure of hydrogen is 400 mm. or 300 mm. and 2·6 for 200 mm. In the packed bulb the temperature coefficient is 1·39 and independent of pressure.

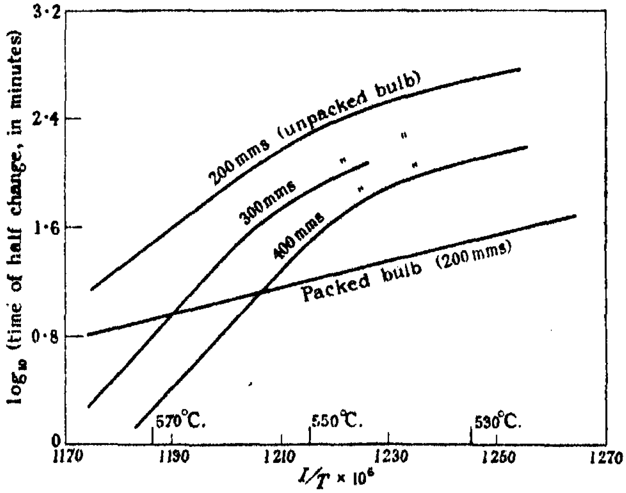


FIG. 2.—Influence of temperature on the rate of combination of oxygen and hydrogen.

Influence of the Pressure of Oxygen and of Hydrogen respectively.

The results so far described show that the behaviour of oxygen-hydrogen mixtures is the same in vessels of porcelain and in vessels of silica. The separate effects of the hydrogen and of the oxygen on the reaction rate have now been determined more accurately than was possible in the silica vessels.

Table II contains the results of experiments made with various proportions of the gases. The data are recorded in detail because they also serve as typical examples of the complete course of the reaction.

Table II.—Combination of Oxygen and Hydrogen at 569° C.
t = time ; x = total change in pressure at time t.

A.—Hydrogen pressure, 300 mm. initially throughout. Varying oxygen pressure.

$[O_2]$ = 100.		$[O_2]$ = 150.		$[O_2]$ = 200.		$[O_2]$ = 250.	
t.	x.	t.	x.	t.	x.	t.	x.
′ ″		′ ″		′ ″		′ ″	
0 45	5	0 34	10	0 27	10	0 22	10
1 15	10	1 8	20	0 51	20	0 36	20
1 43	15	1 49	30	1 17	30	0 53	30
2 12	20	2 27	40	1 44	40	1 9	40
3 14	30	3 13	50	2 16	50	1 31	50
3 46	35	4 50	70	2 47	60	1 54	60
4 24	40	6 0	80	3 30	70	2 20	70
5 11	45	7 27	90	4 21	80	3 2	80
5 46	50	9 30	100	5 31	90	3 56	90
7 15	60	13 5	110	6 55	100	5 9	100
11 28	80			10 49	115	7 23	110
15 45	92					11 25	120

B.—Oxygen pressure, 100 mm. Varying hydrogen pressure.

$[H_2]$ = 200.		$[H_2]$ = 250.		$[H_2]$ = 300.		$[H_2]$ = 350.		$[H_2]$ = 400.	
t.	x.	t.	x.	t.	x.	t.	x.	t.	x.
′ ″		′ ″		′ ″		′ ″		′ ″	
1 50	5	1 9	5	0 45	5	0 55	10	0 30	10
3 55	10	2 11	10	1 18	10	1 18	15	0 45	15
6 12	15	3 16	15	1 53	15	1 40	20	0 59	20
8 52	20	4 31	20	2 38	20	2 5	25	1 11	25
11 25	25	5 52	25	3 15	25	2 53	35	1 26	30
14 20	30	7 26	30	3 54	30	3 19	40	1 40	35
17 31	35	10 49	40	4 43	35	4 15	50	1 51	40
21 0	39	13 0	45	6 26	45	5 15	60	2 8	45
		15 27	50	8 8	55	7 54	80	2 21	50
		18 12	55	10 25	65	9 46	90	2 49	60
		24 5	65	12 53	75			3 25	70
				14 15	80			4 25	85
								5 7	90
								7 15	95
								10 0	96·5

From the results in Table II curves were plotted and tangents drawn giving the initial rate of reaction. The relative rates for the various pressures are tabulated below, the smallest value in each series being taken as unity.

Table III.—Combination of Hydrogen and Oxygen at 569° C.

$[H_2]$ = 300 mm.		$[O_2]$ = 100 mm.	
$[O_2]$.	Relative rate of reaction.	$[H_2]$.	Relative rate of reaction.
mm.		mm.	
250	3·67	400	8·37
200	2·38	350	4·88
150	1·71	300	3·02
100	1·00	250	1·84
		200	1·00

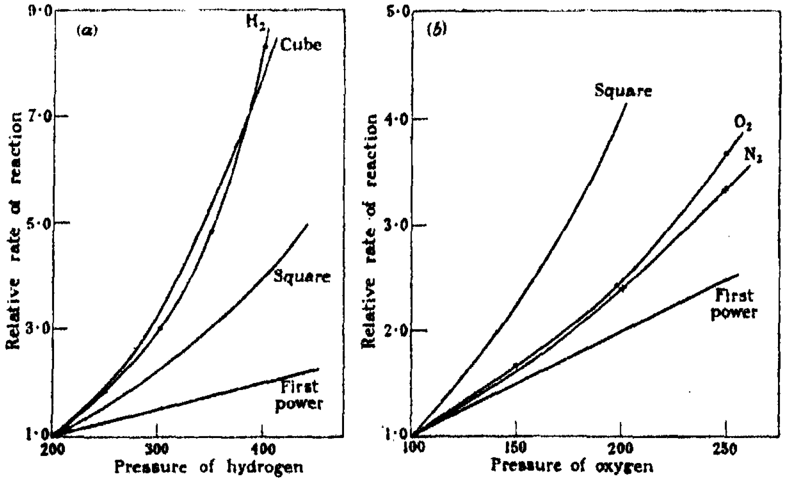


Fig. 3.—Influence of initial pressure of hydrogen and of oxygen on the rate of combination of hydrogen and oxygen at 569° C. (*a*) Constant oxygen pressure of 100 mm. The lines marked "First power," "Square," and "Cube," show how the rate would vary if it were proportional to the respective power of the hydrogen concentration. (*b*) Constant hydrogen pressure of 300 mm. The line marked N_2 shows the effect of a total pressure of nitrogen equal to the *excess* of the oxygen pressure over that corresponding to the rate unity.

Influence of Steam.

The reaction is accelerated by the presence of steam. The influence of steam previously introduced into the vessel is shown by the results summarised in Table IV.

Table IV.—Influence of Steam on Rate of Combination of Oxygen and of Hydrogen at 569° C.

$[H_2] = 200$ mm.; $[O_2] = 100$ mm.

Initial pressure of steam.	Initial rate of reaction in millimetres per second.	Relative rate.
0	0·039	1·00
50	0·053	1·4
100	0·141	3·6
200	0·243	6·2

Influence of other Gases on the Rate of Reaction.

The dependence of the rate of reaction upon the pressure of hydrogen, oxygen and steam cannot be adequately discussed until the effect of various " inert " gases on the reaction has been considered. Helium, argon and nitrogen all have a most remarkable accelerating influence. The helium was obtained from a cylinder of high-grade helium and used without further treatment; argon was prepared by passing 90 per cent. argon six times over red-hot copper and red-hot magnesium and drying with phosphorus pentoxide; nitrogen was obtained by drying cylinder nitrogen, or, in some experiments, by using purified air instead of oxygen for the reaction.

The accelerating effect is greatest with argon and least with helium. Some typical curves showing the effect of argon and nitrogen are shown in fig. 4.

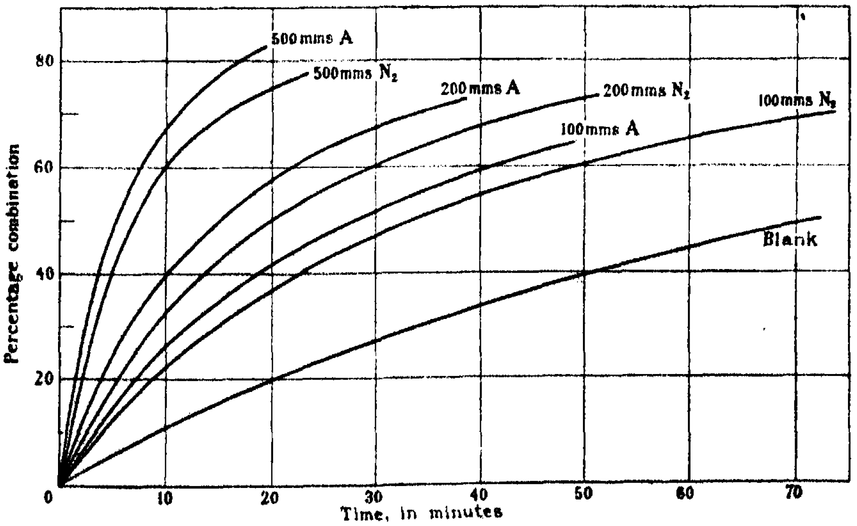


Fig. 4.—Typical curves showing the influence of argon and of nitrogen on the rate of combination of hydrogen and oxygen at 559° C. (The curve marked " blank " refers to 200 mm. H_2 and 100 O_2 without inert gas. The others refer to the same proportions of H_2 and O_2 with the addition of the amount of argon or nitrogen marked on each curve.)

The effect of the inert gas is nearly the same at all stages of the reaction: thus with 200 mm. H_2 and 100 mm. O_2 in the absence of nitrogen the times required for 25 per cent., 50 per cent. and 70 per cent. of the complete reaction are as 1 : 2·8 : 6·1; with 200 mm. N_2 present they are as 1 : 2·7 : 6·1, and with 500 mm. N_2 as 1 : 2·8 : 6·6. This fact facilitates comparison, since the accelerating influence of the inert gas can be represented by a single number which applies to all stages of the reaction and therefore also to the initial rate. The figures given in Table V are averages calculated for each 10 per cent. up to 50 per cent.

The effect of the inert gas increases rather more than linearly with its concentration, but for pressures up to 400 mm. a linear relationship would be sufficiently near the truth.

Table V.—Accelerating Effect of Various Gases on the Combination of Hydrogen and Oxygen.

Temperature.	Added gas.	Pressure of added gas.	Pressure of hydrogen.	Pressure of oxygen.	Rate of reaction relative to rate in absence of inert gas.
569°	Helium	0	300	150	1·00
		300	300	150	2·45
		500	300	150	3·91
569°	Helium	0	200	100	1·00
		300	200	100	1·75
		500	200	100	3·29
569°	Nitrogen	0	200	100	1·00
		50	200	100	1·26
		100	200	100	2·40
		150	200	100	3·33
		200	200	100	4·76
		250	200	100	6·17
559°	Nitrogen	0	200	100	1·00
		100	200	100	2·23
		200	200	100	3·64
		300	200	100	4·61
		500	200	100	10·4
559° (at 569° the reaction is too fast in presence of argon to be conveniently measurable)	Argon	0	200	100	1·00
		100	200	100	2·85
		200	200	100	5·24
		500	200	100	14·2
		700	200	100	19·1

In the previous paper the action of steam was described as autocatalytic; but there is now no reason to assume that its influence is of a kind different from that of the inert gases.

The accelerating influence will be referred to for brevity as the "inert gas effect." Qualitatively it is in the order, steam, argon, nitrogen, helium. A quantitative measure of the relative effect is difficult to give in a concise form because it varies a little with the exact method of computation, and depends for example on whether we compare the acceleration produced by equal amounts of two gases, or compare the quantities which produce the same acceleration. Comparing the acceleration produced by helium with that produced by nitrogen, a ratio of 1 : 3·16 is found when the pressures are 500 mm. and a ratio of 1 : 2·6 when the pressures are 300 mm. The ratios Helium : Argon are 1 : 4·3 and 1 : 4·6 for the same two pressures. Taking all the results into consideration, the following order was arrived at, the ratios being given to the nearest whole number only :—

$$\mathrm{He} : \mathrm{N_2} : \mathrm{A} : \mathrm{H_2O} :: 1 : 3 : 4 : 5$$

Working Hypothesis about the Nature of the Inert Gas Effect.

The pronounced accelerating influence of inert gases on a homogeneous reaction is very unusual, and therefore not improbably operates on that part of the reaction mechanism where the hydrogen-oxygen combination differs from other gas reactions. The reaction is peculiar in that, while proceeding mainly in the gas, it is retarded by the influence of the surface. The only possible explanation of this seems to be that the surface removes, destroys or de-activates something upon which the combination in the gas phase depends, possibly an actual chemical catalyst, but more probably molecules activated by the energy released in the reaction. The reaction chain, which would be propagated by these molecules, may be assumed to be interrupted whenever one of the activated molecules is de-activated by contact with the surface. The inert gas increases the length of time during which any molecule in the gas phase escapes contact with the walls of the vessel, and therefore would lengthen the life of any catalyst or any kind of active molecule which would be destroyed there. We assume for the moment that the inert gas does not itself deprive the active molecules of their energy—if active molecules are in question—but this assumption can be justified on general grounds as will be shown in a later section. Thus the inert gas would be expected to lengthen the reaction chains and increase the rate of reaction. If this view of the matter is correct, the acceleration should be the greater the less readily the inert gas allows other molecules to diffuse through it.

The diffusion coefficient of one gas into another is proportional to the quantity

$$\frac{1}{S_{12}^2}\sqrt{\frac{1}{m_1}+\frac{1}{m_2}},$$

according to the Stefan-Maxwell theory. m_1 and m_2 are the molecular weights of the two gases, and S_{12} is the arithmetic mean of the diameters of the two kinds of molecules. For the present purpose let m_1 refer to the inert gas, and m_2 to the molecules which have to diffuse through it; m_1 is definitely known, m_2 is not. Not only the absolute values of the diffusion coefficient but the relative values also depend upon m_2. If m_2 is large, the molecular weight of the inert gas governs the situation, while if m_2 is small the molecular diameter has most influence. Since we do not yet know m_2, the following table has been constructed, making three assumptions about the nature of the diffusing molecules, namely, that they are H_2, H_2O or H_2O_2. The molecular diameters are taken from the table in Jeans' 'Dynamical Theory of Gases,' except that for hydrogen peroxide, which is taken to be equal to the sum of the diameters of H_2 and O_2. In each series the diffusion coefficient in helium is taken as unity. The diffusion coefficient of oxygen in the different gases will obviously be similar to that of H_2O and H_2O_2, which, it will be seen, are similar to one another.

m_2.	Relative values of $\frac{1}{\text{Diffusion coefficient}}$ (He = 1) for diffusion in					
	H_2.	He.	N_2.	O_2.	A.	H_2O.
$H_2 = 2$	1·0	1·0	2·1	2·0	2·0	2·5
$H_2O = 18$	0·9	1·0	2·8	2·8	2·9	3·0
$H_2O_2 = 34$	0·9	1·0	2·9	2·9	3·05	2·95

From $m_2 = 18$, the order of 1/D for He, N_2, A, H_2O is the same as the order of the accelerations. Steam, argon and nitrogen are rather closer together than they are in respect of their accelerating effect, but the ratio He : N_2 is quantitatively correct. The calculation of the diffusion coefficient is of an approximate nature only, and the assumption that the rate of reaction will vary in a simple inverse ratio with the coefficient is obviously an approximation also, but there seems to be enough parallelism between the two series to support the hypothesis that the inert gas lengthens the reaction chains.

The Inert Gas Effect of Oxygen and Hydrogen.

Oxygen and hydrogen themselves may easily exert an inert gas effect in addition to their normal mass action effect as participants in the reaction, and therefore it becomes clear that the relation between the rate of reaction and the pressure of oxygen and of hydrogen need not be a simple one. From the position of oxygen and hydrogen in the above table we can make an approximate allowance for the inert gas effect they would have if their action were analogous to that of the other gases. That of oxygen should be almost exactly equal to that of nitrogen. Experiment shows the effect of an excess of oxygen on the rate of reaction to be very similar to that of an equal pressure of nitrogen. Thus with 300 mm. H_2 and 250 mm. O_2 the rate of reaction was 2·15 times as great as with 300 mm. H_2 and 150 mm. O_2. With an extra 100 mm. N_2, instead of the extra 100 mm. O_2, added to the equivalent mixture, the rate was 2·40 times as great in one series of experiments and 2·23 times in another. Unless, therefore, there is some reason why oxygen itself should not exert an inert gas effect, the concentration of oxygen must practically disappear from any mass action equation governing the rate of change. If, on the other hand, we assume that the rate of reaction is proportional to the first power of the oxygen concentration, then there is still room for an inert gas effect but less than half as great as that of nitrogen.

The inert gas effect of hydrogen would be much smaller and nearly equal to that of helium; the effect of hydrogen on the rate of reaction, however, is even greater than that of oxygen, so that there is room for a definite mass action effect in addition to the full inert gas effect. Perhaps the simplest way of disentangling the effects is to make various assumptions about the power of the hydrogen concentration which appears in the mass action equation, make allowance for the effect of this on the rate of reaction and see how any residual effect compares with that of helium. This is done in the following table. $[O_2] = 100$ mm. throughout.

$[H_2]$ in excess of equivalent amount (200).	Relative rate of reaction $R_{obs.}$	Relative rate calculated assuming proportionality to			$\frac{R_{obs.}}{R_1}$	$\frac{R_{obs.}}{R_2}$	$\frac{R_{obs.}}{R_3}$	Relative rate for equal excess of helium.
		$[H_2]$ R_1	$[H_2]^2$ R_2	$[H_2]^3$ R_3				
0	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00
50	1·84	1·25	1·56	1·95	1·47	1·18	0·94	1·05
100	3·02	1·50	2·25	3·38	2·01	1·34	0·89	1·12
150	4·88	1·75	3·05	5·35	2·78	1·60	0·91	1·22
200	8·37	2·00	4·00	8·00	4·18	2·09	1·04	1·35

If the mass action effect were proportional to $[H_2]$ it would leave rather too large an amount to be accounted for by the inert gas effect; proportionality to $[H_2]^2$ would leave an inert gas effect more or less comparable with that of helium. On the other hand, it will be observed that we could assume proportionality to $[H_2]^3$ and no inert gas effect.

Mechanism of the Reaction.

In the light of these considerations we may attempt to construct a working hypothesis about the mechanism of the reaction. The number of possible assumptions that can be made about the nature of the reaction chains is rather large; the consequences of many have been worked out and shown to be inconsistent with the facts. The equations are too lengthy to give in full, but the following qualitative arguments will show in a general way the line of investigation leading to the hypothesis finally adopted.

1. We assume that reaction chains are set up by the handing on of the energy of freshly formed steam molecules, and that the chains are broken by the de-activation of molecules.

2. The retarding effect of an increased surface, and the accelerating effect of inert gases are assumed to be accounted for by the fact that the most important kind of de-activation *is that which occurs in a heterogeneous reaction at the walls* of the vessel. No de-activating effect is attributed to the molecules of inert gases.

3. We may further assume that the reaction is stopped principally by the de-activation of molecules of some particular kind X. If X were H_2O_2 or H_2O or anything which did not react with oxygen, the inert gas effect of oxygen should be about the same as that of nitrogen, thus leaving no mass action effect for the oxygen. $[O_2]$ would therefore have to cancel out of any mass action equation. This condition can be satisfied by assuming a specific de-activating influence of oxygen itself at some point in the chain. But this assumption makes the principal way in which the chains are broken depend no longer on the surface, and we thus have to seek another explanation of the inert gas effect. If X is an activated oxygen molecule the rate of reaction on most assumptions becomes proportional to $[O_2]$, but under these circumstances the inert gas effect of oxygen need no longer be normal. The life of active oxygen molecules in the gas phase is increased by the presence of inert gas: the period which elapses before a given oxygen molecule reaches the surface is increased also by the presence of other oxygen molecules. But whereas the inert gas turns back active molecules of oxygen moving towards the surface

without depriving them of their internal energy (assumption 2), other oxygen molecules which they encounter are quite likely to take over the internal energy as well as the component of translational momentum directed towards the surface. Oxygen, therefore, would be much more transparent to the passage of energy than foreign gases. We shall assume, since we cannot have a full inert gas effect and a mass action effect, that the molecules immediately activated by the freshly formed steam are molecules of oxygen.

4. On this hypothesis the inert gas effect of hydrogen will not be normal since the activated oxygen will probably react with the next hydrogen it meets, to give a temporary molecule of H_2O_2, which may fall apart again, or go on to meet another hydrogen.

Thus the inert gas effect of hydrogen might be very small. To account for the observed effect, $[H_2]^3$ would have to appear in the mass action equation.

5. The natural assumption to make now is that the activated oxygen is removed by reaction with adsorbed hydrogen at the surface.

6. The following mechanism, which is the simplest possible, gives a very fair representation of the results which are found at the highest temperatures :—

$$
\begin{aligned}
&(1)\ 2H_2 + O_2 = 2H_2O^* \\
&(2)\ H_2O^* + O_2 = H_2O + O_2^* \\
&(3)\ O_2^* + 2H_2 = 2H_2O^* \\
&(4)\ O_2^* + H_2 = H_2O \text{ (at surface)} \\
&(5)\ H_2O^* \rightarrow H_2O.
\end{aligned}
$$

When the chains are established

$$\frac{d\,[O_2^*]}{dt} = k_2\,[H_2O^*]\,[O_2] - k_3\,[O_2^*]\,[H_2]^2 - k_4\,[O_2^*]\,[H_2] = 0,$$

$$\frac{d\,[H_2O^*]}{dt} = k_1[H_2]^2[O_2] + k_3[O_2^*][H_2]^2 - k_5\,[H_2O^*] - k_2\,[H_2O^*]\,[O_2] = 0.$$

The rate of reaction is governed principally by (3), whence

$$\text{Rate} = \frac{k_1k_2k_3\,[H_2]^3\,[O_2]^2}{k_3k_5\,[H_2] + k_4k_5 + k_2k_4\,[O_2]},$$

It is evident that we can have a high and variable order of reaction. When chains are effectively propagated k_5 is small and the rate reduces to

$$\frac{k_1k_3\,[H_2]^3\,[O_2]}{k_4},$$

---

* Means an activated molecule.

k_4 depends upon the rate at which active oxygen diffuses to the surface, and is therefore inversely proportional to some function of the inert gas concentration.

7. If we assume, as an alternative to the above,

$$
\begin{aligned}
&(1)\ H_2 + O_2 = H_2O_2{}^*\\
&(2)\ H_2O_2{}^* + H_2 = 2H_2O^*\\
&(3)\ H_2O^* + O_2 = H_2O + O_2{}^*\\
&(4)\ H_2 + O_2{}^* = H_2O_2{}^*\\
&(5)\ H_2O_2{}^* \rightarrow H_2O + \tfrac{1}{2}O_2 \text{ (at surface)}\\
&(6)\ H_2O^* \rightarrow H_2O
\end{aligned}
$$

we find, when certain simplifying assumptions about the relative size of the constants are made

$$\text{Rate} = \frac{k_1 k_2 [H_2]^2 [O_2]}{k_5}.$$

k_5 is decreased by inert gases ; oxygen and hydrogen should have the same kind of effect as other gases. This expression therefore agrees with the facts in regard to hydrogen but not in regard to oxygen unless it be assumed that for some reason oxygen does not impede the passage to the walls of the vessel of the energy contained in active H_2O_2 molecules. This may be true, but could not be predicted.

Of the two alternatives the first has the advantage of simplicity and may be taken as our working hypothesis.

8. Since it is probable that several chain mechanisms may operate simultaneously, there is no difficulty in seeing how the total order of reaction can change with temperature and pressure.

9. The question arises whether the chains branch ; that is, when two molecules of steam are formed, whether each originates a new chain or only one of them. The average distance of any molecule from the walls is of the order of 1 cm., that is, about 10^5 free paths. If a chain branched once in every few free paths, much more than the whole quantity of gas present would be involved in one single chain. Thus, either the chains do not branch, or, if they do branch, the fruitfulness of collisions must be small. If some rare kind of ternary collision between $O_2{}^*$ and $2H_2$ is required, we can still account for a very large rate of reaction by assuming branching.

---

* Means an activated molecule.

Alternative Hypothesis about the Action of Inert Gases.

Another possible way in which an inert gas could accelerate the reaction should be mentioned. It could catalyse the combination of hydrogen and oxygen to form hydrogen peroxide, if the molecule formed by the collision of a molecule of hydrogen with one of oxygen were only capable of persisting when the liberated energy could be removed by a simultaneous collision with a third molecule. This is analogous to Herzfeld's theory of the combination of atoms to form diatomic molecules, where he supposes a ternary collision to be necessary. Without the inert gas the H_2O_2 falls apart into H_2 and O_2 again; if de-activated the H_2O_2 decomposes normally to oxygen and water. The most serious objection to this interesting possibility is that it seems to be inconsistent with the existence of reaction chains, and if reaction chains are abandoned the influence of the surface becomes difficult to account for.

Elastic Collisions between Inert Gas Molecules and Activated Molecules.

We have had to assume that the inert gas does not dissipate the energy of activation of the active molecules. The assumption may seem arbitrary, but there are a number of facts which show that collisions between unlike molecules result much less frequently in a transformation of internal energy than might have been supposed. When a mixture of chlorine and ozone is exposed to light practically every quantum of light absorbed by the chlorine ultimately causes the decomposition of an ozone molecule, whatever the ozone concentration may be. Thus the activated chlorine seems to survive an indefinite number of collisions without dissipation of the energy, until at length an ozone molecule capable of being decomposed is encountered.* The decomposition of acetaldehyde is bimolecular, although chemically it might perfectly well be unimolecular ($CH_3CHO = CH_4 + CO$). Single molecules of acetaldehyde can be caused to decompose by collision with hydrogen or nitrogen molecules, but these collisions are found to be only about one-tenth as effective as collisions between pairs of aldehyde molecules.† The specificity of energy transfers is further shown by the fact that hydrogen can communicate activation energy to such molecules as that of dimethyl ether, while other gases apparently cannot.‡ Fluorescence is often extinguished in concentrated solutions, showing, as Perrin points out,§ that the molecules of the fluorescent substance exert effects upon one another, resulting in degradation of internal energy, which the solvent molecules cannot exert upon them.

---

* Weigert and Nicolai, 'Zeitschrift f. physikal. Chem.,' vol. 131, p. 267 (1928).

† 'Roy. Soc. Proc.,' A, vol. 116, p. 163 (1927).

‡ *Ibid.*, vol. 115, p. 215.

§ 'Compt. Rend.,' vol. 184, p. 1097 (1927).

Summary.

In continuation of a previous investigation, the combination of hydrogen and oxygen at temperatures between 500° and 600° C. has been studied by a static method, the reaction vessels being of porcelain instead of silica as before.

It is confirmed that above 500° a homogeneous reaction comes into play. The order is high and somewhat variable ; the temperature coefficient is high, and the reaction at higher temperatures is retarded by increasing the surface exposed to the gases. The rate is approximately proportional to the cube of the hydrogen concentration and to a power of the oxygen concentration which is greater than unity.

The reaction is very markedly accelerated by steam, but also by gases such as helium, nitrogen and argon. Thus the action of the steam is not necessarily autocatalytic in the ordinary sense.

The results can be interpreted by assuming reaction chains to be propagated through the gas. These chains are broken by de-activation of molecules in a heterogeneous reaction at the walls of the vessel, and lengthened by the presence of the inert gases, which increase the time during which molecules escape contact with the walls. The order of effectiveness of the various inert gases can be correlated with the diffusion coefficients. Oxygen and hydrogen themselves may have an " inert gas effect " in addition to their mass-action effect as participants in the reaction : consideration of this throws some light on the nature of the reaction chains. Collisions of activated oxygen molecules with molecules of inert gases appear to be elastic.

We are indebted to the Royal Society and to the Chemical Society for grants with which part of the apparatus was obtained.

On Series in the Spark Spectra of Germanium.

By K. R. Rao, M.A., D.Sc., and Prof. A. L. Narayan, M.A., D.Sc., F.Inst.P.

(Communicated by Lord Rayleigh, F.R.S.—Received April 16, 1928.)

[Plate 11.]

Introductory.

Of the spectra of the elements of Group IV of the Periodic Table, that of germanium has been the subject of only a few investigations, probably due to the minute quantities in which the metal is available. Among the earliest measurements of the spectrum were those made by Rowland and Tatnall, Kobb, Exner and Haschek and Eder and Valenta, which were collected by Kayser.* For many years our knowledge of the structure of the spectrum was limited practically to the few constant frequency interval groups pointed out by Paulson† among the arc lines.

McLennan and McLay‡ identified, with the aid of these groups, a few of the important combination multiplets involving the chief characteristic terms $^3\overline{P}_{012}\,^1D_2\,^1S_0$ and $^3P\,^1P$ of the neutral atom.

Ireton§ has measured the spectrum between λ 2257 and λ 320 under "hot spark" conditions and recorded the important lines. More extensive measurements are those of Carroll‖ below λ 1624, and he has also discovered the chief members of the primary series of trebly ionised germanium.

Notable among these investigations of the spark spectrum is the highly suggestive contribution by Dr. Lunt.¶ He measured, with a four-prism glass spectrograph, the spectrum of the spark in air and a discharge through the vapour of germanium chloride under "different temperatures" and divided the lines into three distinctive groups. But these measurements were limited to the visible region, and no attempt was made to group the observed pairs into any series system. Further, as will be presently seen, some of these pairs are really members of a triplet system.

In continuation of the recent work in this laboratory on the spark spectra of tin,* the analysis of the spark spectra of germanium has been taken up by the authors.

---

* 'Handbuch,' vol. 5.

† 'Acta Univ. Lund.,' vol. 10, No. 12 (1914).

‡ 'Trans. Roy. Soc. Canada,' vol. 3, p. 355 (1926).

§ 'Trans. Roy. Soc. Canada,' vol. 3, p. 103 (1924).

‖ 'Phil. Trans.,' A, vol. 225, p. 357 (1926).

¶ 'M.N.R.A.S.,' vol. 85, p. 38 (1924).

It may also be remarked that observations in the extreme ultra-violet in the case of almost all the elements hitherto studied, are generally confined only to the conditions which generate the spectra of singly and multiply ionised atoms in addition to those of neutral atoms. As many of the important combination lines of the arc spectrum of Ge are likely to occur in this region, it has also been thought that an analysis of the spark spectra of the element would eliminate the more prominent spark lines and perhaps tend to simplify the identification of the chief arc lines.

Experimental.

The material used for the work was pure germanium metal, a small quantity of which was specially prepared for this use by Adam Hilger, Ltd., and was reported to be of a high degree of purity.

In connection with their recent work on the series spectra of ionised tin, the authors have measured with accuracy the spark spectrum of tin. So rods of this metal were used in preference to aluminium to serve as carriers. A small quantity of germanium was fused to the tips of pure tin rods, which were utilised as electrodes between which the spark took place. To obtain the enhanced spectra a large X-ray coil capable of giving a 10-inch spark in air was employed, the primary being fed by a current of 10–15 amps. The secondary contained a battery of condensers of capacity about 0·02 m.fd. and an auxiliary spark in air in series with the experimental gap served to elicit different stages of the spectrum.

The spark was primarily studied in an atmosphere of hydrogen, under varying conditions. The arc *in vacuo* which would have given further valuable information about the spectrum of Ge II was not attempted as the metal available was hardly sufficient for the purpose. To classify the lines and assign them to the several stages recourse was had to the usual method of changing the self-induction or capacity in the circuit, and only those lines were taken into the series systems about which there is no doubt as to the origin or the stage of ionisation.

The spectra were photographed by a 5-foot concave grating and a large Hilger quartz spectrograph. The measurements of the lines of wave-length above λ 2600 were made on the grating spectrum and for other regions the quartz spectrograph is of fairly large dispersion. The region below λ 2400 was photographed on Schumann plates made by Hilger. In measuring up the quartz spectrogram the method of hyperbolic interpolation by the Hartmann formula was used, and generally the measurements may be considered accurate to about 0·05 A.U.

---

* Rao, Sn. IV, 'Phys. Soc. Proc.,' vol. 39, p. 408 (1927), and In II, Ga II (Sn III), *ibid.*, p. 161. Also Narayan and Rao, Sn II, 'Z. f. Physik,' vol. 45, p. 350 (1927).

Notation and Numeration of Terms.

There seems to be now a fairly general agreement as to the notation of spectroscopic terms which is an adaptation of the generalised system introduced by Landé. The multiplicity of the term is indicated by an index number on the upper left of the term symbol, and the inner quantum number by a suffix on the lower right. The terms S, P, D, . . . themselves sufficiently represent the azimuthal number, or more correctly the quantised azimuthal sum denoted by "*l*," while the dashing of terms is utilised to distinguish between terms resulting from different conditions of electron transitions.

The numeration of series terms, however, is differently adopted by different writers. In the present paper as in the previous one on Sn II,* the authors have followed A. Fowler in numbering the deepest term of any class as 1, the next deepest as 2, and so on in the case of terms which form a Rydberg sequence. Other terms of the same type which arise from different electron configurations are differentiated by the use of the letters *a*, *b*, *c*, etc.

Series of Ge II.

Theoretical.—From the position of Ge in the table of elements it must be expected that its first spark spectrum consists of doublet and quartet terms, and that the structure resembles generally the arc spectrum of Ga or the first spark spectrum of any of the elements C, Si, Sn or Pb of the same chemical group. These latter have been almost completely analysed and in each case the deepest term was found to be of ^{2}P type. The present work on the series in Ge II completes our knowledge of the spectral series of the singly ionised elements of this group.

According to the theory of spectra developed by Pauli, Heisenberg and Hund† on the basis of the conclusions of Russell and Saunders‡ from the calcium spectrum, the characteristic terms arising out of any given electron configuration can be predicted with certainty, and it will be seen that the results of the analysis of these spark spectra are in complete accord with theoretical predictions.

---

* *Loc. cit.*

† 'Z. f. Physik,' vol. 33, p. 345 (125). See also Fowler and Hartree, 'Roy. Soc. Proc.,' A, vol. 111, p. 83 (1926), and McLennan, *ibid.*, vol. 112, p. 76 (1926).

‡ 'Astrophys. J.,' vol. 61, p. 38 (1925).

The structure diagram of singly ionised germanium will be written, in the manner suggested by Saha,* as

$$\begin{array}{ccccccc} K &&&&&& \\ 2 &&&&&& \\ & L_1 & L_2 &&&& \\ & 2 & 6 &&&& \\ && M_1 & M_2 & M_3 && \\ && 2 & 6 & 10 && \\ &&& N_1 & N_2 & N_3 & N_4 \\ &&& 2 & 1 && \\ &&&& O_1 & O_2 & \dots \\ &&&&& \dots & \end{array}$$

(1) Since in atoms of this type, where there are three electrons outside the corresponding rare gas shell, only one electron normally is in an uncompleted shell, all the underlying shells being complete and therefore spectrally neutral, the transitions of this electron determines the chief low-lying terms in the emission spectrum. As each electron is supposed to carry with it intrinsically a doublet character, the spectrum consists of ${}^2P, {}^2S, {}^2D, {}^2F \ldots$ terms and their higher Rydberg sequences. These constitute the usual doublet system.

Next, if one or both of the inner "*s*" electrons be excited, new groups of terms arise which should also be found to be characteristic of the spectrum.

(2) The configuration $1\,N_1\,2\,N_2$ (*i.e.*, one "*s*" electron and two "*p*" electrons) gives $a\,{}^4\overline{P}$, $a\,{}^2D$, $a\,{}^2\overline{P}$, $a\,{}^2S$. This important group of terms has been found to be present in all the spectra C II, Si II, Sn II and Pb II.

(3) The configuration $1\,N_1\,1\,N_2\,1\,O_1$ yields $b\,{}^2P$, $b\,{}^4P$, $b\,{}^2P$.

(4) When all the three electrons are in the "*p*" orbit, *i.e.*, $3\,N_2$, the terms are $c\,{}^4\overline{S}$, $c\,{}^2P$, $c\,{}^2\overline{P}$.

(3) and (4) are higher terms of the spectrum and give rise to combination lines in the far ultra-violet region and are difficult to be identified.

The analysis of the spark spectrum of germanium showed the existence of the chief regular doublet terms and of $a\,{}^2D$, $a\,{}^2\overline{P}$ and probably of $a\,{}^4\overline{P}$ of the $1\,N_1\,2\,N_2$ configuration, and is thus in entire agreement with that expected from Hund's theory.

Identification of the Series.—On analogy with the spectrum of Sn II it seemed predominantly clear from the beginning of the investigation that the strong pair λ 6021, λ 5894 with a frequency interval 360 constitutes the chief member of the second principal series of Ge II. The recurrence of this frequency interval among the other lines of Ge II showed that this difference must be characteristic of the energy levels of the spectrum. Further evidence is afforded in support of the identification by the location of the corresponding pair λλ 4037, 3923 in the spectrum of As III, a preliminary report of which has been recently communicated to 'Nature.' Table I showed the relativity-doublet-law variations in Ga-like atoms.

---

* 'Phil. Mag.,' vol. 3, p. 1265 (1927), and vol. 4, p. 193 (1927).

Table I.

	$1\,^2S_1—2\,^2P_1$.	Diff.	$1\,^2S_1—2\,^2P_2$.	Diff.	$\Delta\nu$ $2\,^2P_1—2\,^2P_2$.	$\sqrt[4]{\Delta\nu}$.
Ga I	[8266]		[8374]		—	
		8338		8589		
Ge II	16604		16963		359·6	4·36
		8160		8524		
As III	24764		25487		723·2	5·19

Details of the various doublets which are found definitely to belong to the first spark spectrum of Ge are given in Table II. The first two of these and partially the third were recorded previously by Lunt, the remaining identifications being new. The first column indicates the series designation which is given to these lines in the present paper. The limit $2\,^2P_1$ calculated from the pairs of the sharp series is 50426 and agrees with the directly obtained value 49632.

The elimination of the strong lines with separation 360 in the visible region leaves the outstanding pair λλ 5179, 5132 which is strong and diffuse and can be no other than the fundamental doublet $1\,^2D—1\,^2F$. There is no indication of any satellite attached to the lines, and the "F" level, as in the case of tin, is unresolved.

A characteristic feature in the spectra of Si II, Sn II, Pb II is the presence of a strong combination series of the type $a\,^2D_{3,\,2}—m\,^2F$, of which the first member was unmistakable in each case; it appears even in the under water arc spectrum of tin and lead. The variation of the $a\,^2D_3$ term in these elements indicates that it may have a value of about 64000 for Ge II (4 × 16000).

$a\,^2D_3$ (ν/4).	
C II	30432
Si II	19125
Ge II	[15863]
Sn II	14555
Pb II	11834

Table II.—Doublet System of Ge II.

Classification.	λ I.A. (int.).	(vac.). inf.	Δν.
1 $^2S_1$—2 $^2P_1$	6021·14 (8)	16603·6	359·6
1 $^2S_1$—2 $^2P_2$	5893·46 (10)	16963·2	
2 $^2P_1$—2 $^2S_1$	6336·31 (4)	15777·7	360·2
2 $^2P_2$—2 $^2S_1$	6484·32 (5)	15417·5	
a $^2D_2$—2 $^2P_2$	6966·55 W	14350·4	169·4
a $^2D_3$—2 $^2P_2$	7049·74 S	14181·0	359·5
a $^2D_2$—2 $^2P_1$	7145·54 W	13990·9	
2 $^2P_2$—2 $^2D_2$	4824·56 (4)	20721·5	42·0
2 $^2P_2$—2 $^2D_3$	4814·80 (10)	20763·5	360·9
2 $^2P_1$—2 $^2D_2$	4741·96 (9)	21082·4	
1 $^2D_2$—1 ^2F	5131·70 (10)	19481·3	176·4
1 $^2D_3$—1 ^2F	5178·60 (10)	19304·9	
a $^2D_2$—1 ^2F	2845·42 (9)	35133·9	169·0
a $^2D_3$—1 ^2F	2831·80 (9)	35302·9	

(The wave-lengths of the combination a ^2D—2 ^2P are adopted from Lunt.)

The approximate position of the pair a $^2D_{3,\,2}$—1 ^2F will then be about ν 36000. Examination of the spectrum at once revealed the intense doublet λλ 2845, 2832 included in Table II. Even the measurements of Eder and Valenta, who record only very few of the Ge spark lines, contained this strong pair.

The frequency difference 169 occurs again in the combination pair a ^2D—2 ^2P shown in the above Table.

An interesting point which must be noted is that this difference a $^2D_2$—a $^2D_3$ in all these elements is only slightly less than the corresponding difference 1 $^2D_2$—1 $^2D_3$. For

	a $^2D_2$—a $^2D_3$.	1 $^2D_2$—1 $^2D_3$.
C II	—	—
Si II	15·91	16·61
Ge II	169·0	176·4
Sn II	625·5	647·4
Pb II	780	776·7

The location of the members of the primary series of the spectrum is next considered. In such an attempt it is necessary to get some idea of the magnitude of the doublet separation of the lowest ^2P levels. To determine this by extrapolation from a Rydberg formula applied to the members of the secondary series will only be a rough approximation. A separation of about 1800 can be inferred from analogy of the spectra of the elements of the same chemical group. A further clue is afforded by the spectrum of Ge I. The simplest transition in the arc spectrum of Ge, when the outermost electron jumps from the O_1 to the N_2 orbit, results in the emission of the characteristic $^3P\ ^3\bar{P}$ group of six strong lines. When the electron is completely removed there remains the singly charged Ge-ion characterised by the basic ^2P term. From the observations made by Hund,* it appears that while $^3\bar{P}_0$ and $^3\bar{P}_1$ of Ge I converge to $^2P_1$ of Ge II, $^3\bar{P}_2$ and $^1\bar{P}_1$ must converge to $^2P_2$. That is, $^2P_1 - {}^2P_2$ must be given by the difference $\infty\ ^3\bar{P}_1 - \infty\ ^3\bar{P}_2$. This limiting difference cannot be evaluated as a series of $^3P\ ^3\bar{P}$ groups are not yet known. But $1\ ^3\bar{P}_1 - 1\ ^3\bar{P}_2$ is equal to 1416 and the limiting difference, as pointed out by McLennan,† must certainly be greater than this. So a frequency difference larger than 1416 may be searched for among Ge II lines.

The values of the terms $1\ ^2S_1$ and $1\ ^2D_{2,\ 3}$ can be calculated from the connected system of pairs that have been already located (Table II). So the correctness of any probable pair believed to be $1\ ^2P_{1,\ 2} - 1\ ^2S_1$ can be checked by the observation of the other pairs $1\ ^2P - 1\ ^2D$ and $1\ ^2P - a\ ^2D$, whose exact position becomes calculable through the $1\ ^2P$ term. This method has led to the following series members and the agreement between the calculated and observed lines is so perfectly satisfactory that it lends support to the whole scheme.

---

* 'Z. f. Physik,' vol. 34, p. 296 (1925).

† 'Trans. Roy. Soc. Canada,' vol. 3, p. 355 (1926).

Table III.—Doublet System of Ge II.

	$1\,{}^2P_1$ 128635	*1764*	$1\,{}^2P_2$ 126871
$1\,{}^2S_1$ 66235·7	1602·6 (6) 62399	*1764*	1649·2* 60635
$1\,{}^2D_3$ 47624·9	-		1261·9 (6) 79246 obs. 79246 calc.
			182
$1\,{}^2D_2$ 47801·9	1237·1 (6) 80834 obs. 80833 calc.	*1770*	1264·8 (4) 79064 obs. 79069 calc.
$a\,{}^2D_3$ 63453·9			1577·0 (4) 63412 obs. 63417 calc.
			169
$a\,{}^2D_2$ 63622·9	1538·2 (2) 65011 obs. 65012 calc.	*1768*	1581·2 (0) 63243 obs. 63248 calc.
$2\,{}^2S_1$ 33855	94780 calc.	*1764*	? 93041 obs. 93016 calc.
$2\,{}^2D_3$ 28509			98377 obs. 98362 calc.
$2\,{}^2D_2$ 28551	? 100110 obs. 100084 calc.		98320 calc.

* Excepting this line, which is taken from Ireton's data, all the wave-lengths are from Carroll's measurements.

Other combinations involving the difference 1764 are given in Table IV and the probable classification is indicated against each. Of these the ${}^2P\,{}^2\overline{P}$ group appears to be most probable. The relative intensities are in correct order, and the group is in the appropriate position for, according to Hund, a set of terms resulting from a given configuration of electrons will generally be of the same order of magnitude, the differences in energy value between themselves being due only to a difference in the orientation of the corresponding orbits. However, a reliable check on the correctness of the identification is not possible, as the combination lines between these and other known levels lie far out in the infra-red.

Table IV.—Other Unclassified Pairs. $1\,{}^2P$—X.

Term value.	$1\,{}^2P_1$ 128635		$1\,{}^2P_2$ 126871	Probable classification.
X = 53936	1338·7 (1) 74699	*1775*	1371·3 (2) 72924	$a\,{}^4P_1$
X = 37610	1098·6 (5) 91025	*1763*	1120·3 (2) 89262	$a\,{}^2\overline{P}_1$
X = 34818	1065·9 (1) 93817	*1778*	1086·5 (5) 92039	$a\,{}^2\overline{P}_2$
X = 36698	1087·7 (1) 91937	*1774*	1109·1 (1) 90163	$a\,{}^2S_1$

Term Values.—Table V gives the term values derived through a connected system of members by assuming $1\,{}^2F$ to be equal to 28320 cm.$^{-1}$ which is believed to be not much in error.

Table V.—Term Values of Ge II.

Electron configuration.	Terms calculated.	Terms observed.	Term value.	Effective quantum number.
$2\,N_1\,1\,N_2$	$1\,{}^2P_1$ $1\,{}^2P_2$	$1\,{}^2P_1$ $1\,{}^2P_2$	128635 126871	1·8467 1·8595
$2\,N_1\,1\,O_1$	$1\,{}^2S_1$	$1\,{}^2S_1$	66235·7	2·5736
$2\,N_1\,1\,N_3$	$1\,{}^2D_2$ $1\,{}^2D_3$	$1\,{}^2D_2$ $1\,{}^2D_3$	47801·9 47624·9	3·0295 3·0351
$2\,N_1\,1\,O_2$	$2\,{}^2P_1$ $2\,{}^2P_2$	$2\,{}^2P_1$ $2\,{}^2P_2$	49632·0 49272·5	2·9731 2·9839
$2\,N_1\,1\,P_1$	$2\,{}^2S_1$	$2\,{}^2S_1$	33855·0	3·5998
$2\,N_1\,1\,N_4$	$1\,{}^2F_3$ $1\,{}^2F_4$	$\}\;1\,{}^2F$	[28320]	3·9359
$2\,N_1\,1\,O_3$	$2\,{}^2D_2$ $2\,{}^2D_3$	$2\,{}^2D_2$ $2\,{}^2D_3$	28551·0 28509·0	3·9199 3·9228
$1\,N_1\,2\,N_2$	$a\,{}^4\overline{P}_1$ $a\,{}^4\overline{P}_2$ $a\,{}^4\overline{P}_3$			
	$a\,{}^2D_2$ $a\,{}^2D_3$	$a\,{}^2D_2$ $a\,{}^2D_3$	63622·9 63453·9	2·6259 2·6294
	$a\,{}^2\overline{P}_1$ $a\,{}^2\overline{P}_2$	$a\,{}^2\overline{P}_1$ $a\,{}^2\overline{P}_2$		

In Table VI a comparison is made of the term values of Ga I, Ge II (and As III) and also of the elements of the same vertical group. The terms of As III are derived by applying a Rydberg formula to the first two members of the secondary sharp series, reported in 'Nature,' and therefore must be considered as only approximate.

Table VI.—Comparison of Terms.

Element.	$1\,^2P_1$.	$2\,^2P_1$.	$1\,^2S_1$.	$2\,^2S_1$.	$1\,^2D_3$.	$2\,^2D_3$.	$a\,^2D_2$.	$1\,^2F$.
Ga I	48380	[15326]	23592	10795	13592	7569	—	—
Ge II	32159	12408	16559	8464	11906	7127	15906	[7080]
As III		10987	13739	7495		7182		
C II	196670	64934	80121	39424	51108	28535	100165	27680
Si II............	131818	50632	66323	33852	52466	30800	76514	28265
Ge II	128635	49632	66236	33855	47625	28509	63623	[28320]
Sn II	117679	46160	60767	[33000]	45640	27303	58848	28401
Pb II	121256	46809	61818	32260	52277	28763	55116	28726

In the first part of the table the terms are divided by 1, 4, 9 as usual.

The largest term $1\,^2P_1$ is equal to 128635 giving the second ionisation potential of germanium a value of about 15·88 volts, while the resonance potential is 7·70 volts corresponding to λ 1602·6 ($1\,^2P_1$—$1\,^2S_1$).

Series of Ge III.

Theoretical.—The Hund theory predicts that the different configurations of a two-valence electron system such as that of Ge III will be characterised by terms as follows :—

(*a*) $2\,N_1$ (both the electrons being in an "*s*" orbit) gives only $1\,^1S_0$, the other terms being excluded by Pauli's principle.

$1\,N_1\,1\,O_1$, $1\,N_1\,1\,P_1$ . . . will give higher Rydberg sequences of this $^1S_0$ and a sequence of $^3S_1$ terms.

(*b*) $1\,N_1\,1\,N_2$: the terms obtained are $1\,^1P$, $1\,^3P$.

(*c*) $1\,N_1\,1\,N_3$: terms obtained are $1\,^1D$, $1\,^3D$, and so on.

These form the singlet and triplet system of terms which are derived normally when the "series electron" is shifted through the successive shells.

The anomalous terms are derived, as pointed out by Kitchlu and Saha in the case of Zn I, by keeping the fixed electron in the next higher orbit N_2 and moving the running electron through the successive levels N_2, O_1, etc. Of these terms the group $a\,^3\bar{P}_{012}$, $a\,^1D_2$, $a\,^1S_0$ arising from $2\,N_2$ configuration are the most important as the combination group $1\,{}^3P - a\,{}^3\overline{P}$ is conspicuous in most of these two electron systems.

Of the series of terms thus theoretically indicated the triplets have been almost completely identified in the present work, and these will be dealt with first.

Triplet System.—The first clue to the identification of this system in germanium was the detection of the following series of triplets (Table VII) in the visible and quartz region of the spark spectrum. The first two of these belong to the principal and sharp secondary series respectively, and the last form the inverted and direct members of a diffuse series. The limits calculated from these are :—

$2\,{}^2P_2$ (sharp series) = 93617.

$2\,{}^2P_2$ (diffuse series) = 93998.

They agree closely between themselves and the term $2\,{}^2P_1$ is of the order which may be expected for Ge III.

Table VII.—Triplet System of Ge III.

	$2\,{}^3P_2$ 91037·7		$2\,{}^3P_1$ 91496·9		$2\,{}^3P_0$ 91665·4
$1\,{}^3S_1$ 114959·8	4179·06 (10) 23922·1	*459·2*	4260·85 (8) 23462·9	*168·5*	4291·67 (5) 23294·4
$2\,{}^3S_1$ 62386·6	3489·27 (6) 28651·1	*458·5*	3434·31 (4) 29109·6	*168·7*	3414·53 (3) 29278·3
$1\,{}^3D_3$ 110507·7	5134·70 (10) 19470·0 *106·3*				
$1\,{}^3D_2$ 110614·2	5106·80 (3) 19576·3	*458·7*	5229·34 (6) 19117·6 *70·2*		
$1\,{}^3D_1$ 110685	—		5210·20 (4) 19187·8	*169·4*	5256·61 (3) 19018·4
$2\,{}^3D_3$ 60328	3255·36 (8) 30709·7 *44·2*				
$2\,{}^3D_2$ 60372·2	3260·06 (4) 30665·5 *30·5*	*460·1*	3211·87 (7) 31125·6 *30·6*		
$2\,{}^3D_1$ 60402·8	3263·30 (0) 30635·0		3215·03 (5) 31095·0	*169·2*	3197·62 (5) 31264·2

Table VII (continued).

	$1\,^3D_3$ 110507·7		$1\,^3D_2$ 110614·2		$1\,^3D_1$ 110685·0
$1\,^3F_4$ [63000]	2104·25 (9) 47507·7 *63·3*				
$1\,^3F_3$ 63063·3	2107·06 (4) 47444·4	*106·5*	2102·34 (8) 47550·9 *17·4*		
$1\,^3F_2$ 63080·7	—		2103·11 (2) 47533·5	*70·8*	2099·98 (8) 47604·3

The frequency interval between the first D terms recurs in the strong multiplet, shown above, of the first fundamental series, the position of which is approximately known from the corresponding multiplets of Zn I and Ga II. A similar group of which the chief lines are λλ 3177, 3138, 3044 is also clearly evident in the spectrum of Pb III which is now under investigation by us.

The above triplets form a connected system of members from which the term values are evaluated by taking $1\,^3F_4 = 63000$, which will not be in error by more than a few hundred units.

The identification of the chief series is now made possible. The table of the spark lines of Ge measured by Carroll was scrutinised for the differences $1\,^3D_1—1\,^3D_2 = 70{\cdot}2$ and $1\,^3D_2—1\,^3D_3 = 106{\cdot}3$ for the identification of the six line diffuse multiplet. A further very valuable clue to the detection of these members is afforded by the application of the relativity laws to Zn-like atoms. A satisfactory triplet $1\,^3P—1\,^3S$ had been located previously by one of the authors* and this facilitated extrapolation to Ge III.

Regular doublet sequence :—

Atomic number (Z) and element.	$\Delta\nu$ $1\,^3P_1—1\,^3P_2$	$4\sqrt{\Delta\nu/0{\cdot}0456}$.	Screening constant.	Internal ratio.
30 Zn I	388·9	9·61	20·39	2·05
31 Ga II	932	11·96	19·04	2·07
32 Ge III	[1642]	[13·77]	[18·13]	[2·15]

---

* *Loc. cit.*

Irregular doublet sequence :—

Atomic number (Z) and element.	ν $1\,{}^3P_2$—$1\,{}^3S_1$	$\nu' = \nu - \frac{9R}{25\times16}(Z-A)^2 = 2468\,(Z-A)^2$	Difference.
30 Zn I	20782	18313	26016
31 Ga II	54201	44329	27897
32 Ge III	[94438]	[72226]	

In applying the latter sequence the method of Millikan and Bowen† has been adopted. The method gave the approximate position of $1\,{}^3P_2$—$1\,{}^3S_1$ as 93000 with the difference $1\,{}^3P_1$—$1\,{}^3P_2 = 1600$.

With the aid of these relations the following triplets have been found among Carroll's measurements. The location of the complete six line multiplet exactly in the calculated position is a strong evidence to the correctness of the identification. Table VIII gives the details.

Table VIII.—Triplet System of Ge III (continued).

	$1\,{}^3P_2$ 209398		$1\,{}^3P_1$ 211040		$1\,{}^3P_0$ 211803
$1\,{}^3S_1$ 114959·8	1058·9 (6) 94438	*1642*	1040·8 (5) 96080	*763*	1032·6 (4) 96843
$1\,{}^3D_3$ 110507·7	1011·2 (9) 98892 obs. 98890 calc. *107*				
$1\,{}^3D_2$ 110614·2	1012·3 (6) 98785 obs. 98784 calc. *78*	*1647*	995·7 (8) 100432 obs. 100426 calc. *81*		
$1\,{}^3D_1$ 110685·0	1013·1 (1) 98707 obs. 98713 calc.	*1644*	996·5 (8) 100351 obs. 100355 calc.	*761*	989·0 (6) 101112 obs. 101118 calc.
$2\,{}^3S_1$*	680·2 147011 calc.		672·7 148653 calc.		669·2 149416 calc.
* $2\,{}^3D_3$ 60328	670·8 149070 calc.				
$2\,{}^3D_2$ 60372·2	671·1 149026 calc.		663·7 150668 calc.		
$2\,{}^3D_1$ 60402·8	671·2 148995 calc.		663·8 150637 calc.		660·5 151400 calc.

* *Cf.* Addendum. This triplet as observed by Lunt is λλ 680·59, 673·07, 669·58.

† 'Phys. Rev.,' vol. 26, p. 313 (1925).

Table IX gives other triplets which exhibit the ^3P differences and the classification is indicated where probable.

Table IX.—Other Unclassified Members of Ge III.

Term value.	$1\,{}^3P_2$ 209398		$1\,{}^3P_1$ 211040		$1\,{}^3P_0$ 211803	Possible classification.
137652	1393·8 (8) 71746	*1649*	1362·5 (4) 73395			
124896	1183·4 (8) 84502	*1638*	1160·9 (8) 86140	*771*	1150·6 (8) 86911	$a\,{}^3\overline{P}_1$
123169	1159·7 (8) 86229	*1644*	1138·0 (8) 87873	*764*	1128·2 (3) 88637	
118900	1105·0 (5) 90498	*1634*	1085·4 (5) 92132			?$a\,{}^3\overline{P}_2$
115791	1068·3 (5) 93607	*1640*	1049·9 (5) 95247			$a\,{}^1D_2$
109816	1004·2 (6) 99582	*1643*	987·9 (5) 101225	*753*	980·6 (1) 101978	

The strong triplet ν 84502, 86140, 86911 is approximately in the region where the $^3P\,{}^3\overline{P}$ group may be expected, and these lines may result in the transition $1\,{}^3P_{0,\,12}$ to $a\,{}^3\overline{P}_1$. It is found, however, difficult to select the remaining three lines of the group as there is more than one possible pair in the vicinity of the triplet. Such pairs are included in Table IX. Further, the corresponding group has not been determined in the spectrum of Ga II.

Singlet System.—The singlet system of lines is generally the most difficult to work out. Certainly, recourse must be taken to the measurement of the Zeeman effect of at least a few of the lines before any attempt can be made to determine unambiguously the complete single system. In the present work only the chief singlets $1\,{}^1S_0$—$1\,{}^1P_1$ and $1\,{}^1S_0$—$1\,{}^3P_1$ have been identified, by a comparative study of the variation of the position of these lines in similarly constituted spectra as set out in Table X.

Table X.

	$1\,^1S_0 \overset{\nu}{—} 1\,^3P_1$	Difference.	$1\,^1S_0 \overset{\nu}{—} 1\,^1P_1$	Difference.
Zn I	32502		46745	23955
Ga II			70700	21186
Ge III	[57677]		91886	
Cd I	30656	12694	43692	19344
In II	43350*	11841	63036*	16872
Sn III	55191†		79908†	
Hg I	39413	12977	54065	
Tl II	52390‡	11997	71195 ?	
Pb III	64387*			

* From unpublished results of the authors.
† Green and Loring, ' Phys. Rev.,' vol. 30, p. 585 (1927) (Sn III).
‡ Reported to ' Nature ' by the authors recently.

$1\,^1S_0$—$1\,^1P_1$ of Ga II has already been suggested by Carroll,* so the prominent line from the measurements of Ireton and also of Carroll, λ 1088·3 is evidently $1\,^1S_0$—$1\,^1P_1$. Further, from an examination of the corresponding lines in the spectra of Ge III and Sn III (Table XI), it is seen that there is generally a displacement towards the red with increase of atomic number, a general observation found to be true in several cases.

Table XI.

Element.	$1\,^3P_2$—$1\,^3S_1$	$1\,^3S_1$—$2\,^3P_2$	$1\,^3P_2$—$1\,^3D_3$	$1\,^3D_2$—$2\,^3P_1$	$1\,^3D_3$—$1\,^3F_4$	$1\,^1S_0$—$1\,^1P_1$	$1\,^1S_0$—$1\,^3P_1$
Ge III	94438	23922	98892	19118	47508	91886	[57677]
Sn III*	80405	21801	82607	18690	37602	79908	55191

* Green and Loring, ' Phys. Rev.,' vol. 30, p. 574 (1924).

It is therefore expected that $1\,^1S_0$—$1\,^3P_1$ of Ge III has a frequency greater than 55191. The only strong line in the probable region is λ 1733·8, which is hence considered to be the required resonance line.

---

* ' Phys. Rev.,' vol. 24, p. 205 (1924).

Term Values.—Table XII gives the values of the several terms found through the frequencies of the connected system of triplets. The largest term $1\,{}^1S_0$ is 268717, if the identification of the resonance line is correct, and leads to the third ionisation potential of germanium as being 33·17 volts.

Comparison of Terms.—A comparison between the terms of Zn I, Ga II (*cf.* Addendum) and Ge III is given in Table XIII. The terms have been divided by 1, 4, 9 respectively.

Table XII.—Term Values of Ge III.

Electron configuration.	Terms.	Term values.	Effective quantum number.
$2\,N_1$	$1\,{}^1S_0$	[268717]	1·9166
$1\,N_1\;1\,N_2$	$1\,{}^1P_1$	[176831]	2·3627
	$1\,{}^3P_0$	211803	2·1588
	$1\,{}^3P_1$	211040	2·1627
	$1\,{}^3P_2$	209398	2·1714
$1\,N_1\;1\,O_1$	$2\,{}^1S_0$		
	$1\,{}^3S_1$	114959·8	2·9303
$1\,N_1\;1\,N_3$	$1\,{}^1D$		
	$1\,{}^3D_3$	110507·7	2·9887
	$1\,{}^3D_2$	110614·2	2·9873
	$1\,{}^3D_1$	110685·0	2·9863
$1\,N_1\;1\,O_2$	$2\,{}^1P_1$		
	$2\,{}^3P_0$	91665·4	3·2815
	$2\,{}^3P_1$	91496·9	3·2846
	$2\,{}^3P_2$	91037·7	3·2928
$1\,N_1\;1\,P_1$	$2\,{}^3S_1$	62386·6	3·9777
$1\,N_1\;1\,N_4$	$1\,{}^1F$		
	$1\,{}^3F_4$	[63000]	3·9583
	$1\,{}^3F_3$	63063·3	3·9563
	$1\,{}^3F_2$	63080·7	3·9558
$1\,N_1\;1\,O_3$	$2\,{}^1D$		
	$2\,{}^3D_3$	60328·0	4·0450
	$2\,{}^3D_2$	60372·2	4·0436
	$2\,{}^3D_1$	60402·8	4·0425

Table XIII.

Element.	$1\,{}^3P_2$.	$2\,{}^3P_2$.	$1\,{}^3S_1$.	$2\,{}^3S_1$.	$1\,{}^3D_3$.	$2\,{}^3D_3$.	$1\,{}^3F_4$.	$1\,{}^1S_0$.	$1\,{}^1P_1$.
Zn I	42876	14436	22094	10334	12989	7184	6931	75767	29022
Ga II	29106	—	15558	—	12824	—	[6960]	—	—
Ge III	23266	10115	12773	6932	12279	6703	[7000]	29857	19648

Series of Ge IV.

Theoretical.—In the atom of Ge IV we have a one-electron system which normally gives the simplest type of spectrum. As the electron runs successively through the N_1, N_2, O_1, N_3 . . . shells, the terms $1\,^2S$, $1\,^2P$, $2\,^2S$, $1\,^2D$, etc., are obtained, the largest being $1\,^2S_1$.

The more complicated scheme of doublets and quartets, which result when one or more of the inner group of 10 M_3 electrons is excited, will not be discussed here as none of these terms is found to be in the present work. These terms have not been identified in any of the Cu-like atoms except in the arc spectrum of copper itself.

Identification of the Series.—Of the regular doublet system the members of the chief series occurring in the vacuum grating region were already located by Carroll and this paper contains therefore only the members of the secondary series which may be expected in the region of longer wave-lengths. As in the case of Sn IV, the method adopted in the location of these pairs is to apply the relativity laws to the series of Cu-like atoms. The spectrum of Ga III identified previously by one of the authors has considerably facilitated the present work. Table XIV gives the regular and irregular doublet law progressions in this series of atoms, and it is seen that the members of Ge IV that have been identified are in good agreement in these sequences. Further, the predicted and observed positions of the inverted pair $1\,^2D—2\,^2P$ agree perfectly within observational error.

Table XIV.

	$2\,^2S_1—2\,^2P_2$.	Diff.	$2\,^2S_1—2\,^2P_1$.	Diff.	$2\,^2P_2—2\,^2D_3$.	Diff.	$\Delta\nu$ $2\,^2P_1—2\,^2P_2$.	$\sqrt[4]{\Delta\nu}$.
Cu I	[6246]		[6200]		[6008]		—	
		6928		6629		10374		
Zn II	13174		12829		16382		245·4	3·96
		7374		7184		11564		
Ga III	20556		20018		27946		539·0	4·82
		7572		7173		11374		
Ge IV	28128		27191		39320		937	5·53

Details of the doublets are set forth in Table XV and the term values, which are determined through the value $2\,^2S_1 = 169{,}432$ given by Carroll, have been entered in the same table, which is self-explanatory.

Table XV.—Doublet System* of Ge IV.

Classification.	λ I.A. (int.).	ν (vac.).	Δν.	Term values.
$2\,^2S_1—2\,^2P_2$	3554·19 (8)	28127·8	936·9	
$2\,^2S_1—2\,^2P_1$	3676·65 (6)	27190·9		$2\,^2S_1$ = 169432 $2\,^2P_1$ = 142241 $2\,^2P_2$ = 141204
$2\,^2P_2—2\,^2D_2$	2547·64 (2)	39240·2	80·1	
$2\,^2P_2—2\,^2D_3$	2542·45 (8)	39320·3	936·6	$2\,^2D_2$ = 102064
$2\,^2P_1—2\,^2D_2$	2488·25 (8)	40176·8		$2\,^2D_3$ = 101984
$1\,^2D_2—2\,^2P_1$	2788·36 (8)	35852·8	937·5	
$1\,^2D_3—2\,^2P_2$	2736·01 (8)	36538·8	251·5	
$1\,^2D_2—2\,^2P_2$	2717·30 (2)	36790·3		
$2\,^2P_2—3\,^2S_1$	2343·37 (2)	42660·5	937·0	$3\,^2S_1$ = 98043
$2\,^2P_1—3\,^2S_1$	2293·0 (2)	43597·5		

Using Carroll's values, $1\ D_3 = 177840$, $1\ D_2 = 178094$, the calculated position of the inverted doublet is :—

$$1\,^2D_2—2\,^2P_1 = 35853$$
$$1\,^2D_3—2\,^2P_2 = 36536$$
$$1\,^2D_2—2\,^2P_2 = 36790$$

and agrees perfectly with the observed position.

For completeness it may be added that other term values given by Carroll are as follows :—

$$1\,^2S_1 = 368701$$
$$1\,^2P_2 = 284598$$
$$1\,^2P_1 = 287386.$$

A Characteristic Group of Lines.—A highly interesting feature exhibited by germanium, which calls for further investigation, is the presence of a characteristic group of lines which appears more intensely under weaker stimuli. When the excitation is a maximum the lines disappear or are suppressed considerably, but the inclusion of self-induction, or reduction in the capacity of the circuit,

---

* The authors have since become aware of a note in (November 19, 1927) 'Nature,' by Smith, who mentions pairs 2 S—2 P and 1 D—2 P of the above table.

FIG. 1.

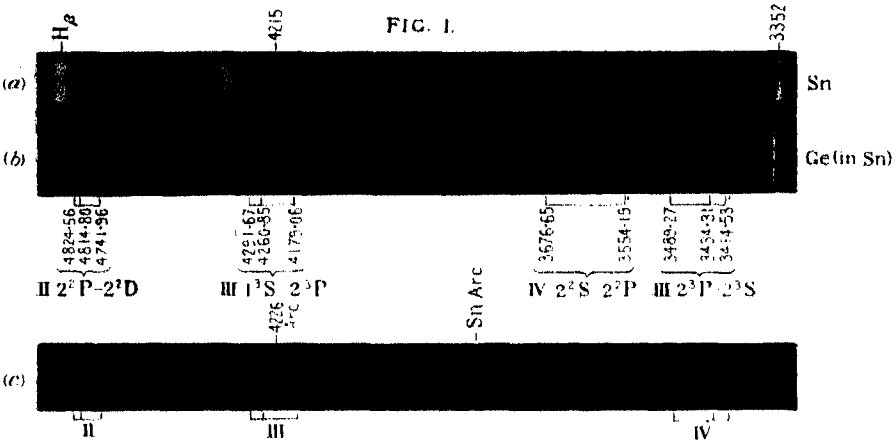


FIG. 2.

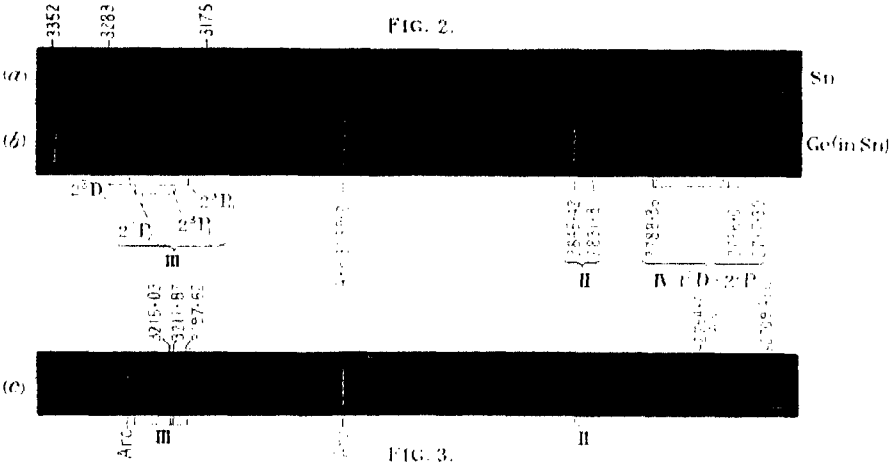


FIG. 3.

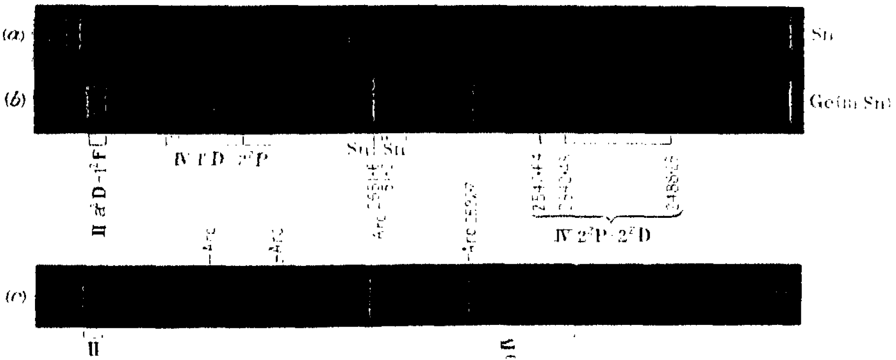
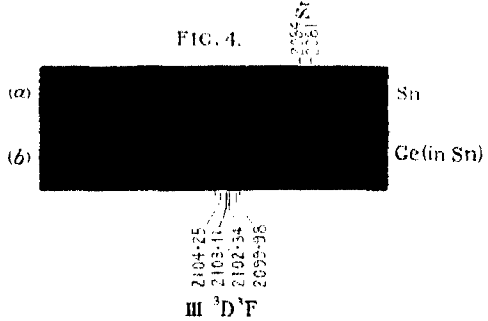


FIG. 4.



brings them out clearly. As such their behaviour is similar to that of arc lines, but none of the previous investigators of the arc spectrum of germanium has recorded the presence of these lines. They have been found to behave in this manner for the first time by Lunt, who worked only in the visible region. The authors have measured many more of such lines in the ultra-violet.

The classification of these and of other lines which form the fainter higher members of the series systems worked out above is still in progress.

Finally, for the sake of convenience all the lines of germanium which have been classified are collected in Table XVI. The wave-lengths below λ 2000 are *in vacuo* and the others in I.A. (air).

Table XVI.—List of Classified Lines.

Wave-length.		(Row.) Lunt.	λ I.A. (int.) Authors.	ν (vac.).	Stage and classification.
Kobb.	Exner and Haschek.				
		7145·8 (W)		13990·9	II $a\,{}^2D_2—2\,{}^2P_1$
		7050·0 (S)		14181·0	II $a\,{}^2D_3—2\,{}^2P_2$
		6966·8 (W)		14350·4	II $a\,{}^2D_2—2\,{}^2P_2$
		6484·4 (5)	6484·32 (6)	15417·5	II $2\,{}^2P_2—2\,{}^2S_1$
6336		6336·4 (3)	6336·31 (4)	15777·7	II $2\,{}^2P_1—2\,{}^2S_1$
6020		6021·25 (10)	6021·14 (8)	16603·6	II $1\,{}^2S_1—2\,{}^2P_1$
5892		5893·6 (10)	5893·46 (10)	16963·2	II $1\,{}^2S_1—2\,{}^2P_2$
5255·5		5256·76 (2)	5256·61 (3)	19018·4	III $1\,{}^3D_1—2\,{}^3P_0$
5228·5		5229·54 (2)	5229·34 (6)	19117·6	III $1\,{}^3D_2—2\,{}^3P_1$
		5210·40 (2)	5210·2 (4)	19187·8	III $1\,{}^3D_1—2\,{}^3P_1$
5177·5		5178·74 (7)	5178·6 (10)	19304·9	II $1\,{}^2D_3—1\,{}^2F$
5134		5134·88 (10)	5134·7 (10)	19470·0	III $1\,{}^3D_3—2\,{}^3P_2$
5131		5131·86 (7)	5131·7 (10)	19481·3	II $1\,{}^2D_2—1\,{}^2F$
		5106·98 (2)	5106·8 (3)	19576·3	III $1\,{}^3D_2—2\,{}^3P_2$
		4824·62 (3)	4824·56 (4)	20721·5	II $2\,{}^2P_2—2\,{}^2D_2$
4813		4814·88 (9)	4814·80 (10)	20763·5	II $2\,{}^2P_2—2\,{}^2D_3$
4742	4744·0 (2)	4742·04 (8)	4741·96 (9)	21082·4	II $2\,{}^2P_1—2\,{}^2D_2$
4291	4291·72 (3)	4291·80 (10)	4291·67 (6)	23294·4	III $1\,{}^3S_1—2\,{}^3P_0$
4260·5	4260·97 (10)	4260·93 (20)	4260·85 (8)	23462·9	III $1\,{}^3S_1—2\,{}^3P_1$
4178	4179·30 (20)	4179·21 (30)	4179·06 (10)	23922·1	III $1\,{}^3S_1—2\,{}^3P_2$
			3676·65 (6)	27190·9	IV $2\,{}^2S_1—2\,{}^2P_1$
			3554·19 (8)	28127·8	IV $2\,{}^2S_1—2\,{}^2P_2$
			3489·27 (6)	28651·1	III $2\,{}^3P_2—2\,{}^3S_1$
			3434·31 (4)	29109·6	III $2\,{}^3P_1—2\,{}^3S_1$
			3414·53 (3)	29278·3	III $2\,{}^3P_0—2\,{}^3S_1$
			3263·30 (0)	30635·0	III $2\,{}^3P_2—2\,{}^3D_1$
			3260·06 (4)	30665·5	III $2\,{}^3P_2—2\,{}^3D_2$
	3255·6 (2)		3255·36 (8)	30709·7	III $2\,{}^3P_2—2\,{}^3D_3$
			3215·03 (5)	31095·0	III $2\,{}^3P_1—2\,{}^3D_1$
			3211·87 (7)	31125·6	III $2\,{}^3P_1—2\,{}^3D_2$
			3197·62 (5)	31264·2	III $2\,{}^3P_0—2\,{}^3D_1$
	2845·5 (4)		2845·42 (10)	35133·9	II $a\,{}^2D_3—1\,{}^2F$
	2831·9 (3)		2831·8 (10)	35302·9	II $a\,{}^2D_2—1\,{}^2F$

Table XVI—(continued).

Wave-length.		(Row.) Lunt.	λ I.A. (int.) Authors.	ν (vac.).	Stage and classification.
Kobb.	Exner and Haschek.				
			2788·36 (8)	35852·8	IV $1\,^2D_2-2\,^2P_1$
			2736·01 (8)	36538·8	IV $1\,^2D_3-2\,^2P_2$
			2717·30 (2)	36790·3	IV $1\,^2D_2-2\,^2P_2$
			2547·64 (2)	39240·2	IV $2\,^2P_2-2\,^2D_3$
	2542·36 (2)		2542·45 (8)	39320·3	IV $2\,^2P_2-2\,^2D_2$
			2488·25 (8)	40176·8	IV $2\,^2P_1-2\,^2D_2$
			2343·37 (2)	42660·5	IV $2\,^2P_2-3\,^2S_1$
			2293·0 (2)	43597·5	IV $2\,^2P_1-3\,^2S_1$
	Ireton. λ I.A. (air).		2107·06 (4)	47444·4	III $1\,^3D_3-1\,^3F_3$
	2104·5 (2)		2104·25 (9)	47507·7	III $1\,^3D_3-1\,^3F_4$
			2103·11 (2)	47533·5	III $1\,^3D_2-1\,^3F_2$
	2102·6 (1)		2102·34 (8)	47550·9	III $1\,^3D_2-1\,^3F_3$
	2100·2 (1)		2099·98 (8)	47604·3	III $1\,^3D_1-1\,^3F_2$
	λ (vac.).		λ (vac.) Carroll.	ν (vac.).	
	1649·2 (3)			60635	II $1\,^2P_2-1\,^2S_1$
	1602·7 (5)		1602·6 (7)	62399	II $1\,^2P_1-1\,^2S_1$
			1581·2 (0)	63243	II $1\,^2P_2-a\,^2D_2$
			1577·0 (4)	63412	II $1\,^2P_2-a\,^2D_3$
			1538·2 (2)	65011	II $1\,^2P_1-a\,^2D_2$
	1500·6 (1)		1500·6 (6)	66640	IV $1\,^2D_3-1\,^2F$
			1494·9 (6)	66894	IV $1\,^2D_2-1\,^2F$
			1264·8 (4)	79064	II $1\,^2P_2-1\,^2D_2$
			1261·9 (6)	79246	II $1\,^2P_2-1\,^2D_3$
	1236·8 (3)		1237·1 (6)	80834	II $1\,^2P_1-1\,^2D_2$
			1229·8 (12)	81314	IV $1\,^2S_1-1\,^2P_1$
	1188·9 (7)		1189·0 (12)	84104	IV $1\,^2S_1-1\,^2P_2$
			1183·4 (8)	84502	III $1\,^3P_2-a\,^3\bar{P}_1$?
	1160·6 (2)		1160·9 (8)	86140	III $1\,^3P_1-a\,^3\bar{P}_1$?
			1150·6 (8)	86911	III $1\,^3P_0-a\,^3\bar{P}_1$?
	1120·2 (2)		1120·3 (2)	89262	II $1\,^2P_2-a\,^2\bar{P}_1$?
			1098·6 (5)	91025	II $1\,^2P_1-a\,^2\bar{P}_1$?
	1088·3 (4)		1088·3 (8)	91886	III $1\,^1S_0-1\,^1P_1$
			1086·5 (5)	92039	II $1\,^2P_2-a\,^2\bar{P}_2$?
			1065·9 (1)	93817	II $1\,^2P_1-a\,^2\bar{P}_2$?
	1058·7 (1)		1058·9 (6)	94438	III $1\,^3P_2-1\,^3S_1$
			1040·8 (5)	96080	III $1\,^3P_1-1\,^3S_1$
			1032–6 (4)	96843	III $1\,^3P_0-1\,^3S_1$
			1013·1 (1)	98707	III $1\,^3P_2-1\,^3D_1$
			1012·3 (6)	98785	III $1\,^3P_2-1\,^3D_2$
			1011·2 (9)	98892	III $1\,^3P_2-1\,^3D_3$
			996·5 (8)	100351	III $1\,^3P_1-1\,^3D_1$
			995·7 (8)	100432	III $1\,^3P_1-1\,^3D_2$
	9890·0 (2)		989·0 (6)	101112	III $1\,^3P_0-1\,^3D_1$
	938·8 (1)		938·9 (6)	106508	IV $1\,^2P_2-1\,^2D_2$
	936·6 (4)		936·7 (9)	106758	IV $1\,^2P_2-1\,^2D_3$
	914·7 (4)		915·0 (8)	109290	IV $1\,^2P_1-1\,^2D_2$
	868·4 (1)		868·3 (6)	115168	IV $1\,^2P_2-2\,^2S_1$
	848·0 (1)		847·8 (6)	117952	IV $1\,^2P_1-2\,^2S_1$

Addendum.

Just when the present paper had been written, the authors received an important paper by Dr. Lang[*] on the Spectra of Gallium, Indium and Germanium. Lang has confirmed the choice of the chief triplets $1\,^3P$—$1\,^3S$ of Ga II and In II made by one of the authors.

The paper contains part of the series system of Ge III worked out in the foregoing pages and there is complete agreement in the identification of the regular triplets. Lang, however, classifies λ 2199·63 as the probable resonance line, but the writers think that this line is too far towards the long wave-length side to be classed as $1\,^1S_0$—$1\,^3P_1$ and does not fit in well in the sequences shown in Tables X and XI.

The terms of Ge III have been calculated by a Rydberg formula applied to the first two triplets of the sharp series, and will certainly be subject to a large error. Lang derives the terms of Ga II by graphical interpolation of $1\,^2P$. One of the writers has identified previously[†] the member $1\,^3D_{1,23}$—$1\,^3F$ of Ga II, and using this member together with those determined by Lang, a reliable term scheme can be obtained by the adoption of a fairly approximate value for $1\,^3F$. The terms so obtained are inserted in Table XIII to make the comparison more interesting.

EXPLANATION OF PLATE 11.

FIG. 1.—(*a*) Spark Spectrum of Sn between λ 4900–3350 (for comparison). (*b*) Spark Spectrum of Germanium in Sn electrodes in the same region without inductance. (*c*) Spark Spectrum of Germanium in Sn electrodes in the same region with inductance.

FIG. 2.—The same spectra in the region λ 3400–2700.

FIG. 3.—The same spectra in the region from λ 3100–2350.

FIG. 4.—(*a*) Strip of Sn Spectrum below 2200 taken with the quartz spectrograph on a Schumann plate, for comparison. (*b*) The corresponding strip of Ge (in Sn) Spectrum showing the fundamental multiplet of Ge III at λ 2103.

The comparison spectra of Sn, figs. 1 (*a*), 2 (*a*), 3 (*a*), 4 (*a*) are all taken with almost the same magnification as those of Ge.

---

[*] 'Phys. Rev.,' vol. 30, p. 762 (1927).

[†] *Loc. cit.*

The Structure of Thin Films. Part XI.—*Oxygenated Derivatives of Benzene.*

By N. K. ADAM.

(*Communicated by Sir William Hardy, F.R.S.—Received April 29, 1928.*)

The object in view when commencing this research was to try and find films of substances containing the benzene nucleus lying flat on the surface of water, in order to measure the cross section of this fundamental unit of Organic Chemistry, parallel to the plane of the ring. The cross section perpendicular to the plane of the benzene ring was measured in Part IV* of this series of papers, and found to be approximately 24 sq. Å.U., a figure in good agreement with the X-ray measurements on naphthalene, anthracene, and benzene.

To achieve this object, it would be necessary not only to know that the benzene nuclei of the molecules chosen were lying flat, but also they must be known to be in contact and not separated from other benzene heads in neighbouring molecules, through some portion of the long hydrocarbon chains attached to the nuclei resting on the surface. To be sure that both these conditions are fulfilled, it would be necessary to find a film in which the area was large, corresponding to the benzene nucleus lying flat; also the area must not alter appreciably with lateral compression, and there must not be any bulky substituents in the nucleus.

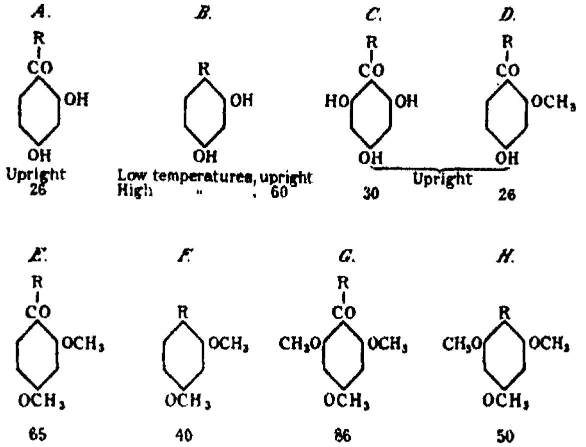
The desired type of film has not been obtained, although more than eight different homologous series have been examined. The substances used were benzene derivatives containing two or more hydroxyl or methoxyl groups in the ring, besides the long chain which is always necessary to give insolubility and stability to the film. The eight principal series used are shown in the following table. R is the long hydrocarbon chain, C_nH_{2n+1}; both alkyl and acyl derivatives have been examined.

In selecting these substances,† it was hoped that the hydroxyl and methoxyl groups in the ring, situated as they are, in some cases, all round the ring, would

---

* 'Roy. Soc. Proc.,' A, vol. 103, p. 676 (1923).

† A great number of the substances have required special preparation though by reactions well known in organic chemistry. I am very much indebted, however, to Prof. P. Karrer, of Zurich, to Sir William Pope and Mr. H. E. F. Notton, and to Messrs. Sharp and Dohme, of Baltimore, for valuable gifts, which have very materially lightened the task of completing this work. Dr. W. A. Berry has also kindly assisted me here during part of the work.



The numbers are the areas at no compression in square Å.U.

attract the water so strongly that the benzene ring would certainly lie down flat. But a complicating factor was not foreseen, namely, that *the same groups which must increase the adhesion to the water also increase the adhesion between adjacent benzene rings standing upright.* Whether the benzene ring lies flat or stands up depends on the relative attractions between the face of the rings and water, and the attractions between adjacent rings placed side by side and upright. If the attraction for the water exceeds that for the next ring in contact face to face, then the ring must lie down; if the reverse is the case, the ring will stand up. It was at first surprising to find that the polyhydroxy compounds (series A to D in the table) had a general tendency to stand up. In fact the only series of this group which formed a film with the benzene ring lying flat were the alkyl resorcinols (B); and these only when the temperature was high, or the hydrocarbon chain short. In most of the polyhydroxy compounds the rings stood upright, the attraction between the adjacent faces of the rings packed upright being great enough to outweigh the attraction of the faces for the water.

Acyl Resorcinols and Phloroglucinols.

Fig. 1 shows the relation between compression on the film and area per molecule, the determinations being made exactly as described previously (Parts VII, VIII, and X*) on members of series A and C. The curves are of the member of series A with 16 carbons in the side chain (palmityl resorcinol) and, in series C, of the 18 carbon member (stearyl phloroglucinol). When compression was first applied, the films contracted steadily, finally reaching the full curve of fig. 1. The dotted curves show approximately the condition of the film when compression was first applied. The films were so rigid that it seems possible the contraction is due to nothing more abstruse than a formation of a film consisting of irregular islands with bare patches of water surface enclosed; after a time under compression, these bare patches are filled up by deformation of the islands. The dotted curves may not therefore be truly "areas per molecule" at all. The full curves show approximately the final areas attained.

---

* 'Roy. Soc. Proc.,' A, vol. 110, p. 423 (1926); vol. 112, p. 362 (1926); and vol. 117, p. 540 (1928).

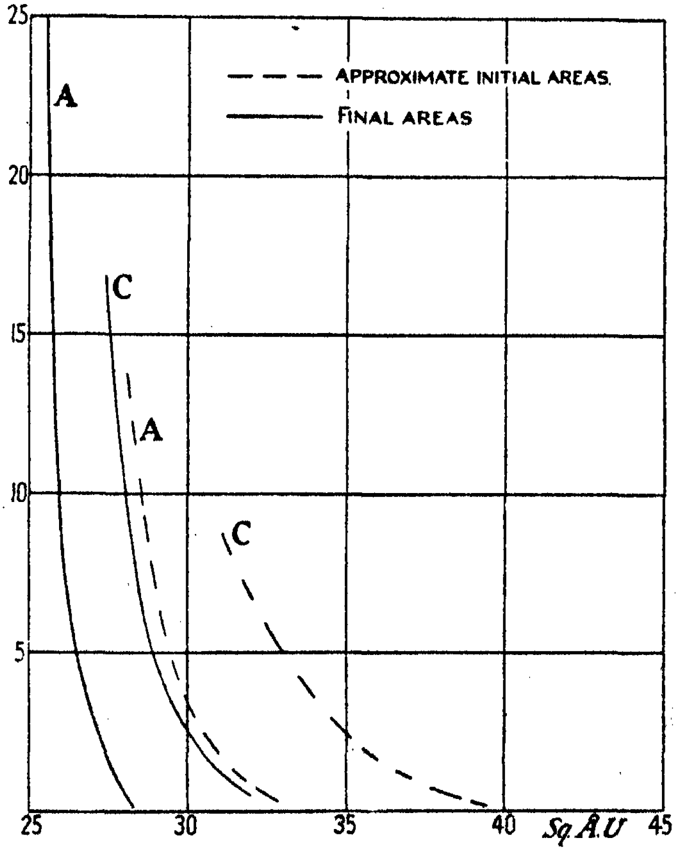


Fig. 1.—Acyl Resorcinols and Phloroglucinols.

..... approximate initial areas. —— final areas.

The final area was, with palmityl and stearyl resorcinols, about 27 sq. A.U. at no compression; with the phloroglucinol derivative it was about 30 sq.

A.U. These areas are probably due to the benzene rings packed closely and nearly, if not quite, vertical in the surface. In hexadecyl phenol, the cross section of the benzene ring with no substituents except the hydroxyl which anchors the molecule to the water was 24 sq. A.U. The difference may well be due to the hydroxyls remote from the water.

Some shorter chain acyl resorcinols were also tried, with from 8 to 12 carbons in the side chains. These gave liquid films, but nevertheless there was some spontaneous contraction. The full details of these films have not been worked out; indeed, somewhat varying results were obtained. As a rule the curve for these substances commenced about 35 sq. A.U., reached 6 dynes per centimetre at 30 sq. A.U. if compressed quickly, and the film contracted slowly. The reason for the somewhat larger areas in these cases is not certain; there is no such sudden change with rise of temperature as is found in the ordinary process of "expansion," but there is a sign of a gradual increase in area, even with the palmityl and stearyl resorcinols, with rising temperature, until at 45° these films were nearly 40 sq. A.U. at no compression.

It appears that, when the chains are at least 16 carbons long the acyl resorcinols form solid films with the benzene rings oriented vertically, or nearly so, at room temperature; at higher temperatures the rings may tilt over somewhat, under low compression, and when first put on, but tend to become vertical under compression and with lapse of time. The shorter chain acyl resorcinols form liquid films resembling the longer chain members at higher temperatures.

Caprylyl phloroglucinol (series C with an 8 carbon side chain) formed a slightly soluble film; accurate measurements were not possible, but it was ascertained that the area at no compression was about 30 sq. A.U., and that the film was liquid. This also probably has the ring vertically oriented. Stearyl phloroglucinol was a solid film even at 41°.

It is clear that both the constitution of the heads and the length of the chains determine the solidity or otherwise of the films. I have recently pointed out that this is a general law in the films*; it is dangerous to generalise about the solidity on the basis of a few observations on only one length of chain.

A compound which probably had the constitution D, with a side chain of 16 carbons, also gave much the same curve as palmityl resorcinol, and therefore had the ring vertical. The compound did not seem of sufficient importance to justify the time necessary to determine the constitution absolutely (the evidence is given at the end of the paper) and it may have had the one methyl group in the para position to the side chain.

---

* 'Trans. Faraday Soc.,' vol. 24, p. 149 (1928).

Alkyl Resorcinols.

The alkyl resorcinols (series B) gave at low temperatures practically the same curve as the preceding, and in this condition they must have the benzene nulcei standing upright. As the temperature is raised, a change takes place, which looks superficially much like the ordinary change from a condensed to an expanded film. Fig. 2 gives the curves for hexadecyl resorcinol. At 1°, the curve is (above 3 dynes per centimetre) definitely that of the benzene ring vertical. At 24° the change to an " expanded " film of area 60 sq. Å.U. at no compression is practically complete. Two members of this series gave reliable results, with 10 and with 16 carbons in the chain. Decyl resorcinol at room temperature gave a curve indistinguishable from the expanded curve of hexadecyl resorcinol at 24°. Evidently, at low temperatures, with a sufficiently

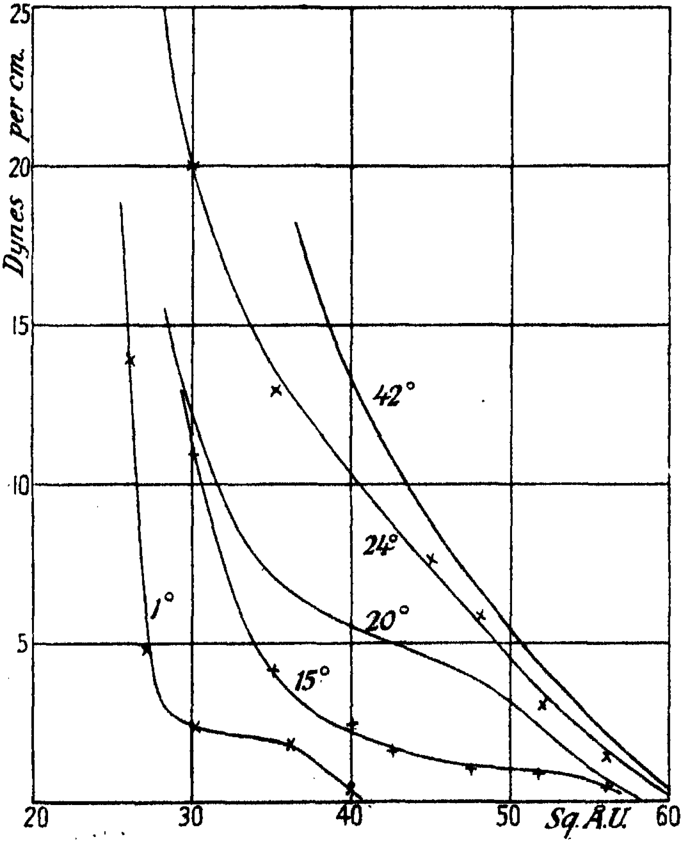


Fig. 2.—B (Hexadecyl Resorcinol).

long chain, the benzene nuclei stand up in the alkyl resorcinols as well as in the acyl resorcinols. The exact nature of the change to the expanded type of film cannot yet be stated, for it is not yet fully understood even with the compounds previously investigated (see Part X). But with most of the previously examined compounds, the heads were so simple that the process of expansion must have depended mainly on the chains. In this case, if the benzene nucleus were to lie down flat, the area would have to become larger, and it is reasonable to suppose that a tendency of the benzene nuclei to lie flat may have something to do with the expansion in this case. The process of expansion may therefore be slightly different in this case from the previously examined cases. Some further discussion of this point will be given later in this paper.

The structure of these expanded films cannot be definitely determined from the areas alone. Very probably the benzene rings lie flat at no compression; it is quite likely, however, that they are not in contact, but that some part of the long chains also lies in direct contact with the water. As the compression is increased, the area decreases so much that the rings must be partly tilted away from the water. Unfortunately, the single measurement of area can never decide both whether there is a tilt and whether there are gaps between adjacent molecules, and we must await the development of independent means of investigation before the exact structure of these films can be determined.

These expanded films are of the "liquid expanded" type. Even decyl resorcinol has a vapour pressure certainly no greater than 0·2 dyne per centimetre—there is no gradual transition into the vapour state of the film.

Methyl Ethers of the Resorcinols and Phloroglucinols.

In the series E to H, which contain methoxyl instead of hydroxyl groups, not a single instance of a condensed film has been found. Fig. 3 gives the compression curves for the dimethyl resorcinol derivatives with 16 carbon side chains (series E and F); fig. 4 gives those for the trimethyl phloroglucinol derivatives (G and H), and fig. 5 the data for the acyl trimethyl phloroglucinol derivatives (G), plotted with the product of compression and area against compression, to show the gradual transition from the liquid to the vapour expanded state as the chains are shortened.

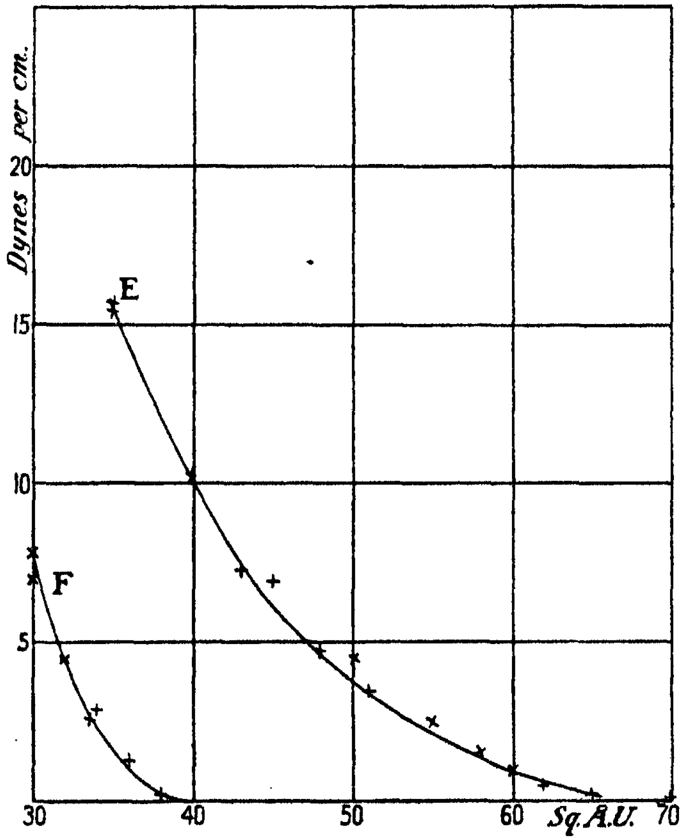


FIG. 3.—Dimethyl Resorcinol Derivatives.

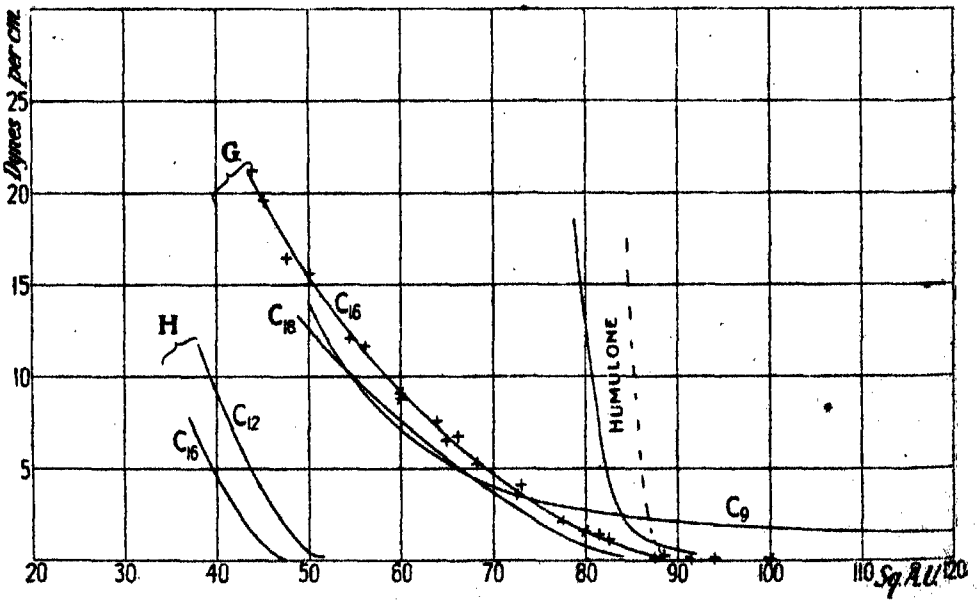


FIG. 4.—Trimethyl Phloroglucinol Derivatives.

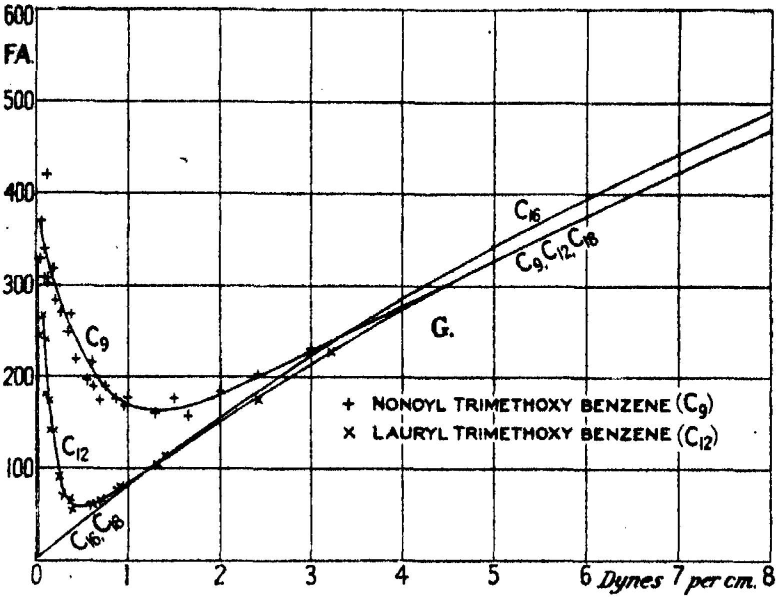


Fig. 5.—Trimethyl Phloroglucinol Derivatives.

The table gives the areas of the actual compounds examined, at no compression, where the films were of the liquid expanded type.

Series.	Compound.	Chain length (carbons).	Area at no compression. Square Å.U. per molecule.	Surface vapour pressure. Dynes per centimetre.
E	Lauryl dimethoxybenzene	12	65	0·07
E	Palmityl dimethoxybenzene	16	66	less than 0·02
F	Hexadecyl dimethoxybenzene	16	40	less than 0·02
G	Nonoyl trimethoxybenzene	9	vapour expanded film	
G	Lauryl trimethoxybenzene	12	vapour expanded film	
G	Palmityl trimethoxybenzene	16	86	0·03
G	Stearyl trimethoxybenzene	18	86	very low
H	Dodecyl trimethoxybenzene	12	51	less than 0·02
H	Hexadecyl trimethoxybenzene	16	48	very low
H	Octadecyl trimethoxybenzene	18	*ca.* 50	very low

Octadecyl trimethoxybenzene was unstable at room temperature, collapsing too fast for accurate measurement. It could be measured at 43°. It is quite common to find that a very long chain makes a film prone to collapse, and rise of temperature normally promotes a tendency to spread, and therefore diminishes the tendency to collapse (except in some cases when the spread film appears to be stable, through some passive resistance to attainment of thermodynamic equilibrium).

The curves are not sensibly different for the different members of the chains, except in series G, when the lateral cohesion between molecules in the expanded film becomes so small when there are only 12 carbons in the side chain that the film is no longer coherent, or liquid expanded. All the curves are of the same type, and much the same as the expanded curve of the alkyl resorcinols. The areas at no compression, however, vary considerably with the constitution of the end group. Thus three methoxyls always impart a greater area to the film than two, and also the carbonyl group in the acyl derivatives is responsible in both di- and tri-methoxybenzene derivatives for a very large increase in area over that of the alkyl derivatives. It would seem that, in these expanded films, both the methoxyl groups and the carbonyl group of the acyl side chain have a tendency to come close to the water.

As with the alkyl resorcinols, it is not possible to say that the area at no compression represents the area of the benzene nucleus with the methoxyls, etc., lying flat on the surface. It is possible that the nuclei are not lying quite flat, and, on the other hand, it is possible that, though flat, they are not in direct contact with each other, being separated by a short length of chain. It is certain that, whatever is the orientation and packing at no compression, this varies continuously as the compression is increased, for the compressibility of the film is so great. Since there must be some self-compression of the film, from the attractive forces between the molecules, there is no means of saying what is the normal, uncompressed state of the film. This well-known difficulty arises in all considerations of the size of molecules, as well as in the problem of film structure.

Cohesional Forces between the Film Molecules.

Although it has not been possible to attain the main object of this research, the measurement of the cross section of benzene parallel to the plane of the ring, still the results give some little information as to the effect of localised adhesional forces—in other words, the chemical affinity of definite groupings of atoms—in orienting the molecules to give the maximum satisfaction of these forces.

We have, first, the unexpected fact that the compounds with several hydroxyl groups in the ring, with only one exception and that only at high temperatures, form films of the condensed type with the benzene nuclei standing upright. The introduction of the two or three hydroxyls must certainly increase the attraction of the face of the ring for the water. But polar or water-soluble groups attract not only each other, but also other polar groups. If the rings were lying flat, then the adhesional affinities of these polar groups would be satisfied on one face only of the ring, since only one face can come into contact with the water. In the vertical orientation, with the faces of the rings packed against the faces of neighbouring rings, there are polar groups on both faces of each ring available to satisfy these affinities; in the flat orientation, the affinities of the polar groups on the upper face of the ring can have available only hydrocarbon chains for their satisfaction, and the attraction between hydrocarbon and polar groups, although not negligible, is less than that between polar groups and polar groups. Therefore we cannot say that the attraction of one face of the benzene ring (with hydroxyl groups) for the water is actually less than the attraction of one face of the ring for the face of another ring,* we can only say that it is not very much greater.

In the compounds with several hydroxyl groups in the ring, the introduction of the carbonyl group greatly increases the tendency of the ring to stand up. It appears, therefore, that the carbonyl group attracts other carbonyl groups, or hydroxyl groups, more than it attracts water, otherwise the tendency of the ring to lie flat would be increased in the same measure as the tendency to pack face to face with neighbouring rings. There is little collateral information as yet as to the adhesional forces or affinities of the carbonyl group, so for the present this must be taken as a somewhat isolated fact. If the carbonyl group did greatly increase the tendency to lie flat, as well as the tendency to pack face to face upright, then one would expect lauryl resorcinol and the lower members of series A to form expanded films, not condensed.

An attempt has been made to increase the attraction of the carbonyl group for the water by forming the films on a solution of hydroxylamine, which reacts with this group, of about 0·95 per cent. strength. The area was only very slightly increased, perhaps not outside the error of observation.

Whether the expanded films in these compounds are exactly the same in structure as those examined in earlier papers it is difficult to say. In most of the earlier types the heads of the molecules were so simple that they would be scarcely capable of much orientation of themselves, or of imparting an orientation to the lower end of the chain; in these compounds probably the chains themselves determine the structure of the expanded films. But in the present series, with the complicated groupings at the ends of the molecules, if the benzene rings lie flat, or tend to do so, the space on the water is so much increased that the chains will naturally tilt over, for in no other way could they pack closely. It seems likely, therefore, that the principal factor determining whether the films are condensed or expanded largely depends on the way in which rings are oriented by the adhesional forces round them. It seems probable that in the "expanded" state of these compounds, the rings are very much tilted over, perhaps completely so, nevertheless the arrangement is an easily disturbed one, as is shown by the great variation of the area with compression. There can be little doubt that the chains are considerably tilted also, arranging themselves so as to fill the space above the heads to the best advantage. Whether there is a definite and uniform tilt of the chains, or a state of vibration which amounts to an average tilt, cannot be decided in these cases, any more than in the cases discussed in Part X.

---

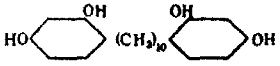
* In the vertical packing of the rings, it is not certain, of course, that each face is packed directly opposite the face of the next ring. It is more likely, by analogy, with the known structure of the naphthalene crystal, that the faces are parallel but echeloned, so that in each row of rings the middle of each face comes opposite the gap between two in the next row. These possibilities as to structure will not seriously affect the argument in this section.

The compounds with methoxyls in the ring all form expanded films; probably therefore the rings lie flat. Now methoxyl has a smaller attraction for water than has hydroxyl, so that the introduction of these groups into the rings should not cause such a large increase in the attraction of the face of the ring for water, as the introduction of hydroxyls in the same positions. The fact that the rings lie flat, or much tilted, shows that the attraction of the face for water is much more increased by the attraction of these methoxyls for the water than is the attraction of the face for the faces of neighbouring rings. Whether this is due to the attraction of a methoxyl group for a neighbouring methoxyl group being small, or to the methoxyl group being of a shape such that the attraction cannot be exercised in full in the packing of the film with upright rings, can scarcely be decided. The fact remains, however, that when the polar groups in the ring are methoxyl instead of hydroxyl, the rings do not pack closely up against each other. Possibly when the rings are flat, the methoxyls can spread out flat also, the oxygen of the methoxyl group coming into very close contact with the water, but if the rings are upright, the methyl groups obstruct the free access of the oxygens on neighbouring molecules to one another.

Fig. 5 shows that the lateral cohesion in the 9 carbon trimethoxybenzene derivative, with the carbonyl group, is quite small; the film is a gaseous one. The 12 carbon member of this series is also a vapour expanded film, and has much less lateral cohesion than the 12 carbon member of any of the other series of this paper, and less than that of the corresponding substance with a CH_2 group in place of a CO group. It is easy to see why this compound should have less lateral adhesion than the hydroxyl compounds, for the methyl groups will clearly diminish the attraction of neighbouring substituent in the nucleus, when the nuclei are lying flat on the surface; the CO group may diminish the lateral adhesion by directing the end of the side chain, next the nucleus, out along the surface, so that the rings are kept somewhat further apart on the surface. Possibly the reason why the acyl trimethoxy derivatives have less lateral adhesion in the films than the acyl dimethoxy derivatives is that the methoxyl groups are spread out flat on the surface; and, when three methoxyls are present, keep the molecules further apart than when there are only two methoxyls.

Compounds with a Resorcinol Group at each end of a Straight Chain.

Mr. Notton has kindly given me some ω ω′ diresorcinyl decane



and its tetra-acetate. These both form gaseous films, the FA—F curves being shown in fig. 6. The cohesional corrections to the perfect gaseous state are fairly considerable in the diresorcinyl decane, and much less in the tetra-acetate;

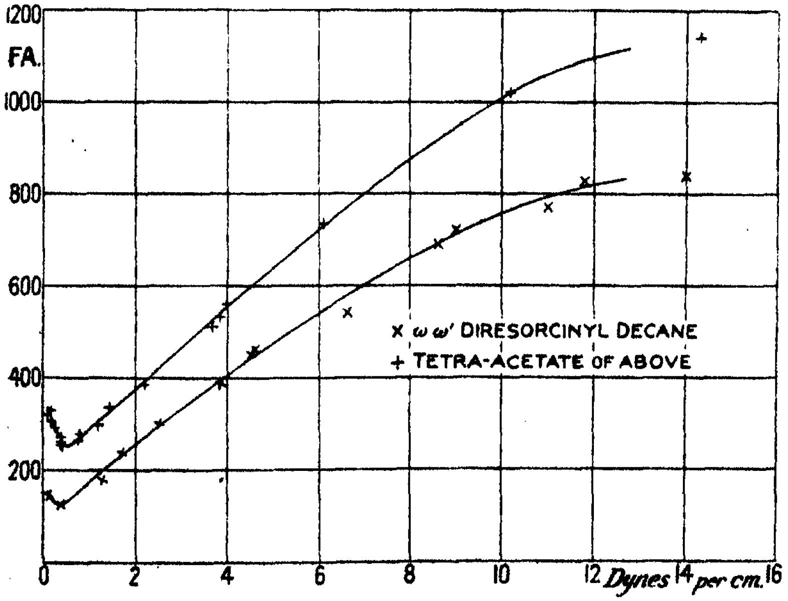
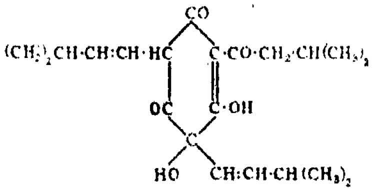


Fig. 6.

a perfectly gaseous film would give a horizontal straight line practically on the line FA = 400, at room temperature. These gaseous films are very much of the type of those previously investigated, in Parts VII to IX, and also of the soluble, adsorbed films of Schofield and Rideal.* There can be no doubt that the molecules are spread out on the surface. It has been shown several times that two definite points of attachment to the water, at widely separated points of the chains, cause the molecules to lie flat and the films to become gaseous, because when the molecules lie flat the opportunities for lateral adhesion are a minimum. The properties of these films are exactly what would be expected from their constitution; the slopes of the curves, above 1 dyne per centimetre, are for the diresorcinyl decane, $d(FA)/dF = 73$, roughly, and for the tetra-acetate, about 90. These slopes, as Schofield and Rideal have pointed out, are the actual areas of the film, per molecule, in the regions where the slopes are constant, as is approximately the case here. It appears that these molecules spread out flat, occupy areas of the order 73 and 90 sq. Å.U. respectively. The tetra-acetyl derivative occupies, as it should, the larger area. These areas are somewhat smaller than those to be expected for these compounds fully extended on the surface, but the method is not one of great accuracy, and it is also possible that the molecules overlap to some extent.

Humulone.

Dr. H. W. Southgate kindly gave me a little of this compound, in the form of its *o*-phenylene diamine compound, from which the humulone itself is easily liberated. The humulone, to which Wieland† has assigned the constitution



with almost complete certainty, gave a film of unique type, the compression curve being shown in fig. 4. It is solid, and only slightly compressible, and has an area at no compression of about 90 sq. Å.U. The dotted curve is what I estimate to be the correct curve for this substance; the actual determinations were slightly complicated by the tendency of the film to collapse, and rather more by the fact that it does not keep well in the free state, even in the dilute solutions in hydrocarbon solvents which are used for applying the substances to the surface. The curves generally obtained approximated to the full curve, but, in one instance, a curve very close to the dotted one was obtained. The interpretation does not differ much within the limits of the errors introduced by this uncertainty as to the true curve.

---

* 'Roy. Soc. Proc.,' vol. 109, p. 57 (1925), and vol. 110, p. 170 (1926).

† 'Ber. Deut. Chem. Ges.,' vol. 58, p. 2013 (1925).

The incompressibility and solidity of the film shows that whatever is the molecular arrangement at no compression, persists when compression is applied, and it seems most likely that the ring is flat on the surface and that adjacent rings are in contact. In fact this compound seems to give a film of exactly that type, which was sought at the beginning of this work as the right type for measuring the benzene ring in its own plane. The area is about 88 sq. Å.U., but unfortunately the substituents in the molecule are so numerous that this figure gives no idea of the true cross section of the benzene nucleus; it is, of course, very much too large, as is known from other data as to the size of this nucleus. It is not even certain that the resistance to compression is caused by the actual contact of the rings; the three chains attached to the ring, if simple, straight chains, should occupy only about 62 sq. Å.U. upright, or packed as usual in condensed films; but the chains are all branched and may in consequence require more space. Moreover, the fact of two substituents being attached to one of the carbon atoms destroys the aromatic character of the nucleus, and this may also have some influence on the size of the nucleus. I incline to the view that the film is packed with the rings, and possibly a short length of the chains, packed in close contact on the surface, the greater part of the chains being packed above the rings, and giving stability to the structure in some way. This view of the structure is, however, far from certain.

Preparation of Material.

The derivatives of trimethoxybenzene were prepared by condensing phloroglucinol trimethyl ether and the distilled acid chloride, in carbon disulphide solution, with freshly sublimed anhydrous ferric chloride, by Kostanecki's method. 32 grams of stearyl chloride, 10 of phloroglucinol trimethyl ether, and 55 c.c. dry carbon disulphide were mixed and 13 grams ferric chloride added in small portions. The reaction proceeded at room temperature fairly gently; after adding all the ferric chloride the mixture was refluxed 1½ hours. Water was added, the carbon disulphide layer separated and evaporated; the solid precipitated by addition of water and by evaporation of carbon disulphide was collected and dissolved in methylated spirit; the free stearic acid was precipitated as barium salt with ammonia and barium chloride; the product was then crystallised from light petroleum. Yield 7 grams, m.p.

67°. Analysis, C = 74·52 per cent., H = 10·67 per cent.; theory 74·59 per cent. and 10·67 per cent. Methoxyl = 21·02 per cent.; theory 21·38 per cent.

Palmityl trimethoxybenzene melted at 54·5°. C = 73·78 per cent., H = 10·55 per cent.; theory 73·85 per cent. and 10·42 per cent.

Lauryl trimethoxybenzene melted at 40°. C = 71·93 per cent., H = 9·90 per cent.; theory 71·94 per cent. and 9·78 per cent.

Nonoyl trimethoxybenzene melted at 34°. C = 69·9 per cent., H = 9·18 per cent.; theory 70·1 per cent. and 9·16 per cent.

In the methoxyl determinations a mixture of 5 volumes hydriodic acid and 4 of acetic anhydride was employed.

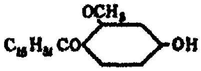
The yields varied considerably. If the carbon disulphide was evaporated straight off from the reaction mixture, the unused acid chloride underwent condensation to the ketone; specimens of laurone and stearone were thus prepared. This reaction of acid chlorides with ferric chloride was described by Hamonet.[*] Methyl alcohol proved useful for re-crystallising the trimethoxybenzene derivatives.

The derivatives of dimethoxybenzene were prepared in a precisely similar manner, using dimethyl resorcinol. The orientation was presumably 1, 2, 4, the acyl group at 1, but this was not proved.

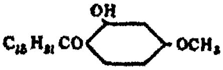
Palmityl dimethoxybenzene melted at 61°. C = 76·53 per cent., H = 10·83 per cent.; theory 76·55 per cent. and 10·72 per cent.

Lauryl dimethoxybenzene melted at 46°. C = 74·65 per cent., H = 10·07 per cent.; theory 74·95 per cent. and 10·07 per cent.

Krafft[†] described a compound melting at 63·5°, formed by the action of aluminium chloride on palmityl chloride and dimethyl resorcin. He presumed this to be palmityl dimethoxybenzene. By this reaction I prepared a compound melting at 66°, which was not identical with the palmityl dimethoxybenzene obtained by using ferric chloride as condensing agent. On methylation with methyl sulphate, however, it passed smoothly into that compound. Presumably my compound melting at 66° is palmityl monomethyl resorcinol, either



or



C = 76·0 per cent., H = 10·7 per cent.; theory 76·2 per cent. and 10·56 per cent. Methoxyl, 9·02 per cent.; theory 8·55 per cent. Aluminium chloride so frequently demethylates the phenol ethers that there can be little doubt that one methyl group has been removed, and the fact that the film was a stable one supports the view that it is the one in the para position to the chain that has been removed, for the films of the para-phenol methyl ethers are not at all stable.

---

[*] 'Bull. Soc. Chim.,' vol. 50, p. 355 (1888).

[†] 'Ber. Deut. Chem. Ges.,' vol. 21, p. 2270 (1888).

The alkyl derivatives of trimethoxy- and dimethoxy-benzene were prepared by reduction of the preceding acyl derivatives by Clemmensen's method, using amalgamated zinc and hydrochloric acid, some methylated spirit being added to promote solution (complete solution was not attained or attempted) and accelerate the reaction. The products were recrystallised from light petroleum or methyl alcohol.

Dodecyl trimethoxybenzene: m.p. 39 to 40°. C = 74·72 per cent., H = 11·05 per cent.; theory 74·98 per cent. and 10·80 per cent.

Hexadecyl trimethoxybenzene: m.p. 54°. C = 76·5 per cent., H = 11·4 per cent.; theory 76·5 per cent. and 11·3 per cent.

Octadecyl trimethoxybenzene: m.p. 62°. C = 77·05 per cent., H = 11·48 per cent.; theory 77·1 per cent. and 11·51 per cent. Methoxyl 21·5 per cent.; theory 22·1 per cent.

Hexadecyl dimethoxybenzene: m.p. 48°. C = 79·4 per cent., H = 11·95 per cent.; theory 79·5 per cent. and 11·70 per cent. Methoxyl 16·8 per cent.; theory 17·1 per cent.

Acyl and Alkyl Resorcinols.

Some of the shorter chain acyl resorcinols, up to 12 carbons, and stearyl resorcinol, were kindly given by Prof. Sir William Pope and Mr. H. E. F. Notton; others, and some alkyl resorcinols, were gifts from Messrs. Sharp and Dohme, of Baltimore.* My best thanks are due for their generosity.

Palmityl and hexadecyl resorcinol were prepared here, by heating together palmitic acid and resorcinol in presence of zinc chloride; after removing excess of palmitic acid by barium chloride, the crude palmityl resorcinol was distilled at 1 mm. pressure, and recrystallised from toluene. The hexadecyl resorcinol was prepared by reduction of the palmityl resorcinol by Clemmensen's method.

Palmityl resorcinol: m.p. 94° to 95°. Hexadecyl resorcinol: m.p. 87°. C = 78·8 per cent., H = 11·6 per cent.; theory 79·0 per cent. and 11·5 per cent.

The stearyl phloroglucinol and caprylyl phloroglucinol (series C) were a gift from Prof. P. Karrer, of Zurich, to whom I am very much indebted.

---

* 'J. Amer. Chem. Soc.,' vol. 48, p. 1688 (1926).

The humulone was liberated from the orthophenylene diamine compound (for which I am indebted to Dr. H. W. Southgate) by treatment with dilute hydrochloric acid and ether or light petroleum, and evaporation of the organic solvent. Being unstable, it was used immediately; the ether was previously distilled over permanganate to remove peroxides, and the results on the specimen prepared with ether were practically identical with those on the other preparation.

Summary.

The surface films of numerous resorcinol and phloroglucinol derivatives, with long side chains, have been examined. The object was to obtain films in which the benzene ring should lie flat on the water surface, closely packed against the rings of neighbouring molecules, in order to measure the cross section of the benzene ring parallel to the plane of the ring.

No films of the required type were formed. In most of the compounds with several hydroxyl groups in the molecule, the rings did not lie flat, but stood up perpendicular to the surface. In those compounds with methoxyl groups the rings probably did lie flat, but the films formed were of the "expanded" type and the packing and structure were very much changed by lateral compression.

The alkyl resorcinols, alone of the eight series examined, formed both condensed and expanded films.

The hydroxyl groups greatly increase the lateral attraction between the rings packed vertically in the films. The carbonyl groups of acyl side chains also appear to increase this lateral attraction, in the polyhydroxy compounds. In the polymethoxy compounds, the carbonyl groups in the side chain appear to come into contact with the water and increase the area.

Compounds with a resorcinol group at both ends of a long chain form gaseous films, the molecules lying flat, in accordance with the previously discovered rule that two widely separated points of attachment to the water on the chain cause the molecules to lie flat and the films to become gaseous. In these gaseous films, the lateral attraction between the flat molecules is diminished by acetylating the hydroxyl groups, and the size of the molecules themselves is increased.

Humulone, one of the "bitter principles" of hops, forms a peculiar film in which the ring is probably flat on the water and closely packed against the neighbouring rings. This compound does not, however, give a measurement of the size of the benzene ring, because so many substituents, and other complications, are present in the molecule.

The Dependance of the Photographic Action of β-Rays on their Velocity.

By C. D. Ellis, Ph.D., Fellow of Trinity College, Cambridge, and
G. H. Aston, B.A., Research Student of Trinity Hall, Cambridge.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 30, 1928.)

The photographic plate has long been used for detecting high speed electrons, but of late several attempts have been made to use it in a more quantitative way for comparing the intensities of two electronic beams. For example, most of the work on β-ray spectra has so far consisted in determining the energies of the various groups of electrons emitted by radioactive bodies by finding the radius of the circle they describe in a uniform magnetic field. This is carried out by placing a photographic plate at some point in their path and noting the point of impact, the only function of the plate being that of a detector. If, however, we wish to compare the number of electrons in two groups of different energies by means of their photographic impression, then we must take into account not only the normal photographic technique of dependance of density on exposure but, in addition, the way the density due to a fixed number of electrons depends on their energy.

It has long been an accepted principle that high speed electrons have less photographic action than slow ones, although we have been unable to find any specific experiments demonstrating this fact. Far less were there any measurements of the exact nature of this dependance, and in some experiments recently carried out in this laboratory on the comparison by a photographic method of the intensities of the electronic groups emitted by radium B and C, Ellis and Wooster* were forced to assume the photographic action ran parallel with the total ionisation per centimetre, that is, varied as $1/\beta^2$. This seemed plausible and was used as a temporary means of obtaining approximate values, but, as was pointed out on that occasion, it was a matter that urgently needed investigation.

We have now made measurements on this point, which show definitely that the photographic action varies on the whole more rapidly than $1/\beta^2$ in the region 200,000 volts onwards. The photographic plate is itself such a complicated instrument that an exact expression of the law of variation has no immediate interest. We have therefore confined ourselves to a special type of plate (Ilford X-ray emulsion), which is always used in this laboratory, and express our results by a curve showing how to reduce photographic densities due to electrons of different speeds to a common basis of true exposure.

---

* 'Roy. Soc. Proc.,' A, vol. 114, p. 276 (1927).

Experimental.

Since we wished to investigate the region of electronic energies above 200,000 volts we were forced to use β-particles of radioactive origin, and this meant inevitably the β-particles of radium B and C, not only because of the ease of obtaining strong sources, but also because it is the only body where the distribution curve of the electrons with energy is known accurately. Gurney* has made a careful investigation by measurement of the charge of how the number of electrons depends on the velocity, and a comparison of his results with the distribution of intensity measured photographically will clearly enable the variation of photographic action with velocity to be determined. The electrons we were using were those forming the continuous spectrum, and the lines when they appeared on the plates were used only for checking up the value of the magnetic field. Our reason for this decision is that from the nature of the experiment we are concerned always with the total number of electrons hitting the plate, that is, even in the presence of a line, with the line plus background. Now in the photographs the lines are apparent in most regions of the spectrum, but in Gurney's electrical method of measuring with its lower resolving power they merge into the background. It is easier to use the continuous spectrum directly and apply a very small correction to Gurney's results for the superposition of non-resolved lines.

Our general procedure may be explained as follows: Two photographs would be taken with different magnetic fields so that at the same position on the plates the electrons hitting had Hρ's of, say, 2000 and 5000. By experiment it was found that to get equal densities the lower velocity photograph had to be exposed one-fifth of the time given to the other, whereas Gurney's measurements showed there were three times as many electrons. The lower velocity electrons are clearly more efficient in affecting the photographic plate, and if we put the photographic action at Hρ 2000 as unity, then that at Hρ 5000 will be 0·6 or three-fifths. This assumes the truth of the "reciprocity law," that the photographic action depends only on the number of electrons hitting the plate and not on how long they take to do it, but this has already been shown to hold by Bothe,† and by Ellis and Wooster.‡ Further, we are making use of the fact that the relative photographic actions of electrons of two specified velocities will be the same at whatever density of the photographic plate the comparison is made. This was shown to be the case by Ellis and Wooster, who found that the density-exposure curves for electrons of different speeds can be made to coincide by taking suitable values for the relative magnitudes of the exposures.

---

* 'Roy. Soc. Proc.,' A, vol. 109, p. 540 (1925).

† 'Z. f. Physik,' vol. 8, p. 243 (1922).

‡ 'Roy. Soc. Proc., A, vol. 114, p. 266 (1927).

We employed the usual type of focussing apparatus used for photographing β-ray spectra, but in order to be justified in using Gurney's results in this way, it was further essential that there should not be any scattered or reflected electrons hitting the plate, or, more accurately, only the same number as in his experiments. There must inevitably be reflection at the source, and from the surrounding metal supports, but this was easy to overcome by making this part of the apparatus identical with Gurney's, and, as in his experiments, using as source a tube filled with radium emanation whose walls were equivalent in stopping power to only 1·5 cm. of air. Electrons reflected elsewhere in the apparatus were prevented from hitting the plate by the usual arrangement of screens, and these appeared to be quite efficient in ensuring a fine beam and reproducing Gurney's conditions. We tested this point carefully by carrying out a second series of experiments when all the slits were narrowed and only a small portion of the plate received any β-rays. The results obtained in this second set agreed well with those of the first.

The experiment was greatly facilitated by a plate-changing apparatus* which enabled five plates to be inserted at once, and then moved up to and removed from the exposure position one by one without releasing the vacuum. In this way three photographs of the region Hρ 2000–2500 would be taken with different exposures. The magnetic field would then be changed to give, say, Hρ 4000–5000, and a fresh plate moved on. It was easy to adjust the exposure so that the density came somewhere in the region covered by the first three plates, so that a simple interpolation gave at once what exposure would have had to have been given in the lower region to obtain exactly the same density as that in the Hρ 4000–5000 region. We could develop 10 plates at once, and this determination of the density-exposure curve was carried out for each development batch by taking three or four plates each time with different exposures in the standard region Hρ 2000–2500. This region was chosen as standard because it is free from strong lines with the exception of Hρ 2256, which served as a check on the magnetic field, and it is near the maximum of the continuous spectrum so that only comparatively short exposures are necessary. In the case of plates exposed in the region of high velocity, comparatively long exposures were necessary to give a sufficient density, and there was then a slight amount of fogging due to the γ-rays. This was very small and easy to correct for by subsidiary experiments.

---

* We would like to take this opportunity of thanking Dr. Wooster for his assistance in the design and construction of this apparatus. It was planned and partly constructed while he was still associated with this work.

We used the ordinary Ilford X-ray emulsion, but our plates were prepared in a special way. A large sheet of patent glass was coated with emulsion and then our plates cut only from the centre portion, where the layer of emulsion was uniform. We are very grateful to Mr. B. V. Storr, of Ilfords, Ltd., for suggesting this procedure and carrying it out for us. These plates were developed in a specially constructed tank, handling 10 plates at once, which was fitted with a plunger whose edges almost scraped the surface of each plate, giving very efficient stirring. Development was carried out for 30 minutes at 18° C. in 1 in 50 Rodinal.

The plates were photometered on a Dobson type photometer giving absolute density differences for each point of the plate. Considerable time was spent on this, each plate being photometered every half millimetre at three different levels.

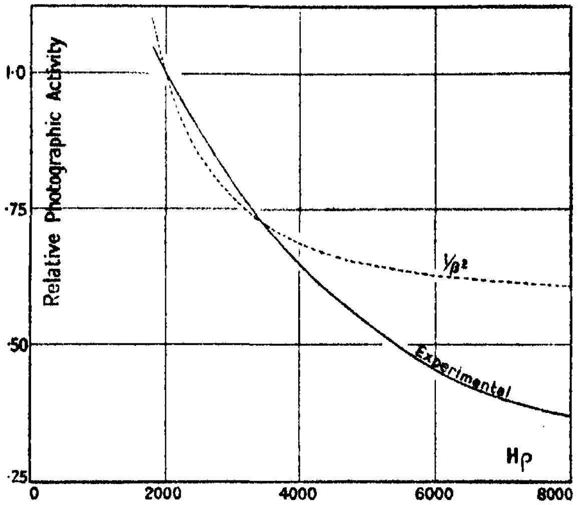
Our final results are shown in Table I, and as a graph in the figure.

Table I.

$H\rho$.	Relative photographic action.
2000	1·00
3000	0·80
4000	0·65
5000	0·54
6000	0·46
7000	0·40
8000	0·36

The accuracy is not very great, partly due to the combination of different measurements used, partly due to the inherent inaccuracies in photographic work. It is clear that any error that may be introduced by using the results of the figure increases with the difference between the values of the momenta of the electrons that are being compared, and we can best express our opinion by stating that the value of for $H\rho$ 4000 may be in error to 5 per cent. but is unlikely to be more than 10 per cent. out.

The dotted curve on the figure shows the variation that would be obtained if the photographic action were to run parallel to the ionisation, that is, vary as $1/\beta^2$. It will be seen that experiment shows initially a less rapid change and, finally, a more rapid change with speed. We think it quite plausible that the photographic action in any small layer will be closely proportional to the ionisation produced there, but that the experimental curve obtained here is due to the superposition of at least two extra disturbing factors. In the first place for speeds below Hρ 4000 an appreciable percentage of the energy is lost in the film, that is, the film must be treated as a thick layer and the loss of velocity and its effect on the ionisation taken into account. In the second place there will be considerable reflection of the β-particles both in the film and at the glass, and the reflected β-particles will be slower and have a greater photographic action. The problem is complicated and experimentally ill-defined, and we have not thought it worth while to attempt a detailed explanation of the experimental curve.



These results enable the measurements of Ellis and Wooster* on the relative intensities of the β-ray groups of radium B and C to be corrected. It is only necessary to take each of the final intensities given in that paper, to divide by β^2 and divide by the appropriate ordinate on the figure. We refrain from publishing a fresh list of these intensities in this paper, partly because it is so simple to effect the transformation and partly because we are checking up these measurements with improved technique and hope to publish a list embodying both the present work and these new measurements.

---

* *Loc. cit.*

We would point out that the curve in the figure renders possible the use of the photographic plate for measurements of intensities of β-ray beams of different velocities strictly only under the same conditions which held in this experiment—that is to say, for the same type of plate and the same type of development. However, general results indicate that the method of development will not affect this curve, but a change of plate might do so. If our suggestion as to the factors controlling the variation with velocity be correct, it will depend on the stopping power of the layer of emulsion and would be different for example with process plates. This is scarcely likely to limit the general use of this curve since X-ray plates would usually be used for this type of work and they have generally the same type of emulsion.

It will be noted that we do not continue our curve to values lower than $H\rho$ 2000. This was due to several causes. The comparison with Gurney's measurements begins here to depend greatly on the exact equality of the stopping power of his emanation tube and ours and, further, we are here approaching the region where the electrons will be completely stopped in the sensitive layer. A great variation might be anticipated between different brands of plates, and in any case this region would be best approached from the low velocity side by more accurate experiments with electrons accelerated from a hot filament.

[*Note added in proof.*—It is not to be expected with these thick layer plates that the curve shown in the figure will continue to increase right down to zero $H\rho$. We have, in fact, recently obtained some evidence that the photographic activity at 600 $H\rho$ is of the order of one-tenth of that at 1400 $H\rho$, indicating a sharp bend over of the curve.]

We have pleasure in expressing our appreciation of the encouragement and help we have received from Prof. Sir Ernest Rutherford, and we acknowledge gratefully a grant to one of us (C.D.E.) from the Government Grant Committee of the Royal Society, which has defrayed some of the costs of the work, and, further, a grant (to G.H.A.) from Trinity Hall, Cambridge, and also one from the Department of Scientific and Industrial Research. We also thank Mr. G. R. Crowe for the preparation of the radioactive sources.

Experiments on the Diffraction of Cathode Rays.—II.

By G. P. Thomson, M.A., Fellow of Corpus Christi College, Cambridge, and Professor of Natural Philosophy in the University of Aberdeen.

(Communicated by Sir Joseph Thomson, F.R.S.—Received May 14, 1928.)

[Plate 12a.]

§ 1. The following paper is an account of further experiments which have been made on the lines of those described in these 'Proceedings' (vol. 117, p. 600), and largely with the same apparatus. A beam of cathode rays passing normally through a very thin film of metal was found to produce a pattern of concentric rings on a photographic plate about 30 cm. away. These could be explained as a diffraction pattern due to the de Broglie waves of the electron, the atoms of the metal crystals being the diffracting system. In the present paper several points of uncertainty are cleared up, and the work extended to other cases.

§ 2. As mentioned in a note at the end of the previous paper, the discrepancy of 6 per cent. between the values of the crystal constants of aluminium and gold determined by X-rays, and those found by applying the de Broglie theory to the diffraction rings formed by the cathode rays, has now been explained. It was due to an error in the measurement of the energy of the cathode rays, and hence their wave-length. The energy was measured by a spark-gap connected in parallel with the discharge tube. In the earlier measurements a considerable length of leads and a rectifying valve were included with the discharge tube, and it now appears that an appreciable fraction of the potential fall occurred in these. When two spark-gaps were used, one connected as before, and the other directly across the discharge, there was 1–2 mm. difference in the readings. The following table shows the values of P, the voltage; D, the diameter of the ring corresponding to reflection from the (2, 0, 0) plane; and $D\sqrt{P}\,(1 + Pe/1200m_0c^2)$, which latter quantity should be constant for any one metal on the de Broglie theory (see previous paper, p. 603). The factor in brackets is the relativity correction, and in the experiments in question never differs from unity by more than 3 per cent.

Film.	P (volts).	D (cm.).	$D\sqrt{P}\,(1 + Pe/1200m_0c^2)$.	"a"
Celluloid	50,000	0·85	195	
	42,500	0·90	189	
	36,000	1·00	193	
	30,500	1·05	186	
	23,200	1·25	193	
	21,000	1·30	190	
	16,800	1·47	191	
	16,100	1·48	189	
	11,500	1·62	175	
	9,800	1·86	185	
		Mean	189	$4 \cdot 17 \times 10^{-8}$
Gold	58,000	1·50	371	
	55,000	1·58	381	
	44,000	1·75	376	
	33,700	2·00	374	
	21,000	2·40	(352)	
		Mean	375·5	$4 \cdot 20 \times 10^{-8}$
Aluminium	64,000	1·47	384	
	57,600	1·62	398	
	45,000	1·78	388	
	34,500	2·00	378	
	21,800	2·42	(361	
		Mean	387	$4 \cdot 07 \times 10^{-8}$
Platinum	45,000	1·85*	402	
	40,000	1·96*	400	
	34,500	2·23*	421	
	25,500	2·46	398	
	18,200	2·75	(375)	
		Mean	405	$3 \cdot 89 \times 10^{-8}$
Aluminium (short camera)	34,500	1·64	310	
	27,500	1·84	310	
	26,200	1·86	305	
	21,800	2·09	312	
		Mean	309	$4 \cdot 00 \times 10^{-8}$
Platinum (short camera)	29,000	1·84†	319	
	24,000	1·98	311	
		Mean .... (weighted)	317	$3 \cdot 89 \times 10^{-8}$

The value of "a" for celluloid is the spacing of the reflecting planes.
The values for the metals are the sides of the unit face-centred cubes of the structure.

* Calculated from the (1, 1, 1) ring by multiplying its diameter by $2/\sqrt{3}$, as the (2, 0, 0) ring was not well defined.

† Mean of four plates with the same potential.

In the first three sets of experiments on metals it will be noticed that the value of $D\sqrt{P}\,(1 + Pe/1200m_0c^2)$ corresponding to the lowest voltage comes low, and has been excluded from the mean. This effect is probably due to the fact that in making the experiments each series was taken with decreasing voltages till the rings became indistinguishable. As the rays were never quite homogeneous, owing to fluctuations of the discharge, it is probable that the faint rings seen with the smallest voltage were chiefly caused by the faster rays, and thus really correspond to a somewhat larger mean voltage than that measured. In the experiments with the short camera (see § 5), two or three Leyden jars were connected across the discharge, and found to have a marked steadying effect, the cathode rays being much more homogeneous when tested by their magnetic deflection. It is noticeable that in these experiments the effect does not occur. The final mean values of "a" compare with those from X-ray measurements as follows :—

Metal.	Al.	Au.	Pt.
Cathode rays	4·035	4·20	$3 \cdot 89 \times 10^{-8}$
X-rays	4·043	4·064	$3 \cdot 913 \times 10^{-8}$

The error is only a little over 1 per cent. on the average, and not systematic. In an accompanying paper Mr. Reid has shown that the results obtained by this method for the size of celluloid rings are in good agreement with those found when the energy of the cathode rays is determined directly from their electric deflection, thus eliminating the possible systematic errors involved in taking spark-gap measurements of the voltage from an induction coil as a measure of the energy of the cathode rays produced.

§ 3. Besides the rings described in the previous paper, rings have been observed for gold with diameters $\sqrt{22 \cdot 7}$, $\sqrt{26 \cdot 6}$ and $\sqrt{35 \cdot 4}$, the diameter of the (2, 2, 0) ring being taken as $\sqrt{8}$. The first two should undoubtedly be $\sqrt{24}$ and $\sqrt{27}$, corresponding to the (4, 2, 2) and (3, 3, 3) plates respectively, the latter being the third order of the (1, 1, 1) plane. The third may be either $\sqrt{35}$ (5, 3, 1) or $\sqrt{36}$, which could be caused by either (2, 0, 0) in the third order, or (4, 4, 2).

The following results were obtained for the relative sizes of the aluminium rings in these experiments, and those to be described in § 5. In all cases the first ring is taken as $\sqrt{4}$, the ring corresponding to (1, 1, 1) of size $\sqrt{3}$ not being found with aluminium. This is in accordance with the results of Konobejewski* for X-ray analysis of rolled aluminium. He found that the thinnest specimens had the axes (1, 0, 0) of their crystals normal to the sheet, thus no (1, 1, 1) planes would come at sufficiently small angles to the beam to give reflections.

---

* 'Z. f. Physik,' vol. 43, p. 741 (1927).

Ring.	1.	2.	3.	4.	5.	6.	7.	8.
Diameter as measured	$\sqrt{4}$	$\sqrt{8 \cdot 00}$	$\sqrt{10 \cdot 9}$	$\sqrt{16 \cdot 5}$	$\sqrt{20 \cdot 4}$	—	$\sqrt{27}$	—
	$\sqrt{4}$	$\sqrt{7 \cdot 65}$	$\sqrt{10 \cdot 6}$	$\sqrt{14 \cdot 8}$	$\sqrt{18 \cdot 6}$	—	—	—
	$\sqrt{4}$	$\sqrt{8 \cdot 05}$	$\sqrt{11 \cdot 05}$	$\sqrt{15 \cdot 8}$	$\sqrt{21 \cdot 4}$	—	$\sqrt{28}$	—
	$\sqrt{4}$	$\sqrt{7 \cdot 93}$	$\sqrt{11 \cdot 4}$	$\sqrt{16 \cdot 0}$	$\sqrt{21 \cdot 0}$	—	—	—
	$\sqrt{4}$	$\sqrt{7 \cdot 95}$	$\sqrt{10 \cdot 8}$	$\sqrt{15 \cdot 8}$	$\sqrt{19 \cdot 3}$	—	$\sqrt{26 \cdot 6}$	$\sqrt{35 \cdot 0}$
	$\sqrt{4}$	$\sqrt{7 \cdot 95}$	$\sqrt{11 \cdot 0}$	$\sqrt{15 \cdot 3}$	$\sqrt{20 \cdot 2}$	—	—	—
	$\sqrt{4}$	$\sqrt{8 \cdot 00}$	$\sqrt{10 \cdot 9}$	$\sqrt{15 \cdot 4}$	$\sqrt{19 \cdot 2}$	$\sqrt{23 \cdot 6}$	$\sqrt{27 \cdot 4}$	$\sqrt{33 \cdot 0}$
Probable true diameter	$\sqrt{4}$	$\sqrt{8}$	$\sqrt{11}$	$\sqrt{16}$	$\sqrt{19}$ or $\sqrt{20}$	$\sqrt{24}$	$\sqrt{27}$	$\sqrt{35}$ or $\sqrt{36}$
Corresponding plane	(2, 0, 0)	(2, 2, 0)	(3, 1, 1)	(4, 0, 0)	(3, 3, 1) or (4, 2, 0)	(4, 2, 2)	(5, 1, 1)	(5, 3, 1) or (6, 0, 0) (4, 4, 2)

§ 4. *Platinum.*—As will be seen from the above results, it has now been found possible* to obtain diffraction patterns with platinum. This has been done in three ways. The first that was successful was to form very thin films of platinum by cathodic spluttering on glass, and remove them with hydrochloric acid. These films were too thin to stand alone and always broke on mounting, so they had to be floated on to a backing of the thinnest celluloid film, and the composite film used. These composite films gave the rings to be expected of platinum, but they were not very sharp. In particular, the rings corresponding to (1, 1, 1) and (2, 0, 0) planes (diameters $\sqrt{3}$ and $\sqrt{4}$) were not resolved. This may be caused by the confusion produced by the second ring due to celluloid, which is about the same diameter, but it is possible that the spluttered platinum is in such small crystals that the resolving power of each is insufficient. With this arrangement rings corresponding to the (2, 2, 0) plane ($\sqrt{8}$) and to the (3, 1, 1) or (2, 2, 2) planes ($\sqrt{11}$ or $\sqrt{12}$) were observed. The rings were continuous, and equally intense all round. The films were thin enough for the filament and bulb of an electric lamp to be clearly visible through them. The next method used was to thin down a piece of foil by hammering, and then gradually reduce it still further in aqua regia. The original foil was too thick for this to be very satisfactory, as the film was apt to go into holes in some places while comparatively thick in others. Nevertheless, one film was obtained which gave fairly good results, though probably only a small proportion of its area was of an effective thickness. The diffraction pattern obtained was quite different in appearance from that of the spluttered films, consisting of a number of isolated spots, showing that only a few crystals were actually diffracting. In most cases pairs of spots occurred at opposite ends of a diameter through the centre of the pattern, as is to be expected theoretically. When measured up, the distances between pairs of spots agreed well with the diameters previously found for the rings of the spluttered films. Spots were found corresponding to the planes (1, 1, 1), (2, 0, 0), (2, 2, 0), (3, 1, 1), *i.e.*, at distances to the ratio $\sqrt{3} : \sqrt{4} : \sqrt{8} : \sqrt{11}$. The potentials were measured before the source of error mentioned in § 2 was realised, and so the absolute value of the grating constant came 7 per cent. too low.

---

* 'Nature,' vol. 120, p. 802 (1927).

Eventually, however, I succeeded in obtaining some very thin platinum leaf made by Messrs. Whiley, Ltd. This was uneven in thickness, and by selecting the thinnest parts, which were fairly transparent, I was able to get diffraction rings without any further thinning in acid, such as was required with aluminium and gold. The rings were nearly continuous, but showed distinct traces of spots, thus being intermediate between the last two results. The same film, when used later for the experiments of § 5, showed rings quite continuous and without spots, but varying in intensity from one quadrant to another (see Plate 12A). Probably this was a slightly different part of the film, as it seems unlikely that platinum would change with keeping. Results with this film are given in the table of § 2. A large number of rings have been measured. Some typical examples are shown below, the innermost ring being taken as $\sqrt{3}$.

Ring.	1.	2.	3.	4.	5.	6.	7.	8.	9.
Diameter as measured	$\sqrt{3}$	$\sqrt{4\cdot00}$	$\sqrt{7\cdot95}$	$\sqrt{11\cdot1}$	$\sqrt{16\cdot2}$	$\sqrt{18\cdot8}$	$\sqrt{23\cdot2}$	$\sqrt{26\cdot8}$	$\sqrt{34\cdot3}$
	$\sqrt{3}$	$\sqrt{4\cdot05}$	$\sqrt{8\cdot05}$	$\sqrt{11\cdot4}$	$\sqrt{15\cdot2}$	$\sqrt{18\cdot4}$	$\sqrt{22\cdot5}$	$\sqrt{26\cdot3}$	$\sqrt{34\cdot2}$
	$\sqrt{3}$	$\sqrt{3\cdot98}$	$\sqrt{7\cdot95}$	$\sqrt{10\cdot8}$	$\sqrt{15\cdot2}$	$\sqrt{18\cdot3}$	$\sqrt{23\cdot8}$	$\sqrt{26\cdot4}$	—
Probable true diameter	$\sqrt{3}$	$\sqrt{4}$	$\sqrt{8}$	$\sqrt{11}$	$\sqrt{16}$	$\sqrt{19}$	$\sqrt{24}$	$\sqrt{27}$	$\sqrt{35}$ or $\sqrt{36}$
Corresponding plane	(1, 1, 1)	(2, 0, 0)	(2, 2, 0)	(3, 1, 1)	(4, 0, 0)	(3, 3, 1)	(4, 2, 2)	(3, 3, 3) or (5, 1, 1)	(5, 3, 1) or $\left\{\begin{matrix}(6, 0, 0)\\(4, 4, 2)\end{matrix}\right\}$

In addition the $\sqrt{12}$ ring could be seen as an edge to the $\sqrt{11}$ along those radii where the (1, 1, 1) ring was strong. This is as it should be, as it is the second order of this ring. The ring $\sqrt{20}$, corresponding to (4, 2, 0), was not directly observed, but it would not be separable from the $\sqrt{19}$ ring. The ring $\sqrt{32}$ was only measurable in one case. It is a second order of $\sqrt{8}$, and in all cases the higher orders are faint. With these exceptions, all the rings which should occur with a face-centred cubic lattice have been observed up to $\sqrt{35}$ and on some plates faint traces of larger rings have been seen.

§ 5. *Magnetic Deflection.*—At an early stage it was noticed that the diffracted pattern was deflected approximately with the main beam of cathode rays by a magnetic field, showing that it was due to rays of about the same speed. It seemed worth while to make careful measurements to see if any slight difference could be detected. For this purpose a new camera was fitted to the old apparatus, shorter and of greater diameter, the distance from film to plate being 24·8 cm. (instead of 32·3), and the brass tube of which it was made having an internal diameter of 12·5 cm. The magnetic field was produced by two large coils of 20 turns and 46·5 cm. diameter, arranged with their planes vertical and 25 cm. apart, this being as near the Helmholtz arrangement as could be managed without further alteration of the apparatus. The centres of the coils were 4 cm. above the line of the undeflected beam, and 8·6 cm. in front of the plate. The soft iron shield between the camera and discharge tube was removed. Usually two exposures were taken on each plate, the current in the coils being reversed. The latter was from 5 amps. to 8·5 amps., according to the speed of the rays used.

In the approximately homogeneous field produced by the coils, the whole pattern was shifted about 3 cm. up or down (see Plate 12A). As far as the eye could judge, the main beam remained the centre of the pattern. To measure the possible displacement accurately it was necessary to mark the centre of the main beam. This was done by looking at the plate by reflected light, when the solarised spot due to the main beam could be clearly seen, and its centre marked by removing a minute portion of film with the point of a needle. From this point the distances to various rings with and against the magnetic deflection were measured. This was done for each of the two exposures on the plate. Experiments were made with platinum, using the film described above, and on aluminium using one of the films used in the previous paper.

For aluminium the $\sqrt{3}$ ring is absent, owing to the special arrangements of the crystals in the leaf, and measurements were taken of the $\sqrt{4}$ and $\sqrt{8}$ rings, or rather of spots on them. For platinum the $\sqrt{3}$, $\sqrt{4}$ and $\sqrt{11}$ rings were measured as, on the whole, the clearest (see Plate 12A for reproductions).

The results are shown in the table, the two exposures on each plate having been meaned.

Ring.	Aluminium.		Platinum.		
	$\sqrt{4}$	$\sqrt{8}$	$\sqrt{3}$	$\sqrt{4}$	$\sqrt{11}$
Difference in radii in mm.	+ 0·14	+ 0·3	+ 0·15	− 0·2	—
	+ 0·3	+ 0·22	+ 0·2	+ 0·5	—
	− 0·1	+ 0·05	+ 1·0	+ 0·6	+ 0·6
	—	—	− 0·05	− 0·05	+ 0·05
Mean	+ 0·11	+ 0·19	+ 0·33	+ 0·21	+ 0·32

The plus sign means that the radius drawn from the main beam in the direction of the magnetic deflection was the longer. The final mean is 0·23, or a shift of the rings with regard to the centre of 0·115. The mean magnetic deflection of the centre was 2·8 cm. A calculation shows that the second order effect due to the plate not being normal to the deflected rays, and to the curvature of their path, would give a positive effect corresponding to a mean shift of the centre of 0·09. (The ratio of the difference in radii to the deflection of the centre comes out as $3\phi^2$, where ϕ is the semi-angle of the cone of diffracted rays.) Thus most of the effect is accounted for, and the residuum is definitely less than the probable error. The speed of the diffracted rays is certainly the same as that of the main beam to less than 1 per cent.

In these experiments it was necessary to connect a condenser of two or three Leyden jars across the discharge tube. This reduced greatly the inhomogeneity of the original beam, and consequently the spreading caused by the magnetic field, which otherwise badly blurred the rings. It was found that the parts of the rings towards the original centre were clearer than without the magnetic field, while those on the other side were more blurred. This is an effect due to the rays not being quite homogeneous even with the condenser, and is exactly analogous to the achromatising of Newton's rings by viewing them through a prism.

§ 6. *Resolving Power.*—It is of interest to consider what can be deduced as to the nature of the waves and crystals from an observation of the width of the rings, as shown, for example, by finding whether two rings close together can be seen as separate.

The width of the rings is due to a variety of causes—the finite size of the original beam, the limited extent of the several crystals, and the lack of homogeneity of the rays; but it will also be affected by the question of how many coherent waves form the train associated with each electron. The last two causes are closely connected, and on some views are different aspects of the same effect.

In practice the first cause is probably the largest, at least when a condenser is used, but observations of resolving power will give a lower limit to the size of the crystals, and to the number of waves in a train.

If the beam is considered infinitely fine and the number of elements in the crystal infinite, the resolving power can be calculated in terms of the number of waves as follows: assume a group of waves for which the intensity is constant from $\lambda - \Delta\lambda$ to $\lambda + \Delta\lambda$, and vanishes outside this range. If p is the order of the reflection $p\lambda = 2d \sin\theta$, and if θ_1, θ_2 are the values of θ for two rings which are just resolved, $2p\Delta\lambda = 2d(\theta_2 - \theta_1)\cos\theta = 2d(\theta_2 - \theta_1)$ approximately. The number of waves in the train can be found as follows: suppose they are all in phase at $x = 0$, then at $x = r\lambda$ the displacement will be

$$\int_{-\Delta\lambda}^{\Delta\lambda} \cos\frac{2\pi x}{\lambda + t}\,dt = -\frac{\lambda^2}{2\pi x}\left[\sin 2\pi r\left(1 - \frac{t}{\lambda}\right)\right]_{-\Delta\lambda}^{+\Delta\lambda}$$

if $\Delta\lambda/\lambda$ is small. For r an integer this becomes $\frac{\lambda}{\pi r}\sin 2\pi r\frac{\Delta\lambda}{\lambda}$. Thus the amplitude vanishes when $2\pi r\frac{\Delta\lambda}{\lambda} = \pi$, and $2r$, the total number of waves on both sides of the origin, is $\lambda/\Delta\lambda$. Thus $2r = p\lambda/d\,(\theta_2 - \theta_1) = 2\theta/(\theta_2 - \theta_1)$.

In these experiments the rings $\sqrt{24}$ and $\sqrt{27}$ were easily resolved, $\sqrt{11}$ and $\sqrt{12}$ just barely; but it is safe to say that they would have been fully resolved but for the finite width of the original beam. This gives

$$2r = 2\sqrt{11}/(\sqrt{12} - \sqrt{11}) = 48.$$

In all probability the number was larger, as the width of the beam would by itself account for most of the observed width of the diffraction rings.

§ 7. *Inclined Films.*—A few experiments were made with films inclined to the rays, the angle between the film normal and the rays being about 30°. If the crystals of the film are arranged quite at random, no change should be produced by inclining the film, but if the crystals have dominant directions with regard to the plane of the film, changes will be produced in the relative intensities of different parts of the same ring, and of the different rings as a whole. A silver film (prepared by Mr. Ironside, see p. 668 *infra*) showed a marked difference in the brightness of different parts of the (2, 0, 0) ring when inclined, though when the film was normal to the rays the ring was of uniform intensity. The brightest parts were at the ends of a diameter parallel to the axis about which the film had been turned. This is what would be expected if there was a tendency for the crystals to be set with one axis normal to the film. The only planes that can reflect the rays are those which make a small angle with them, since the wave-length is small compared with the spacing, and the only (2, 0, 0) planes which would fulfil this condition would be those whose normals nearly coincided with the line about which the film had been turned. That all the crystals are not placed in this way is shown by the fact that the (1, 1, 1) plane appears with the film in the usual position. This ring in the inclined film is brightest on the diameter at right angles to that for the (2, 0, 0) ring, and as the two rings are close together the whole has an elliptical appearance, shown in fig. 3 of Plate 12A. For an aluminium crystal which showed in the normal position a number of spots on the (2, 0, 0) ring, a pair on the diameter parallel to the axis of turning of the film showed strongly. A few other spots were, however, visible on the ring, so the orientation of the crystals with their axes perpendicular to the film is not complete (see § 3).

Discussion.

The results of this paper confirm those of the former in being in agreement with the de Broglie theory of electron waves. The complicated diffraction patterns observed were in all cases in complete agreement with those predicted by the theory and the known structure of the metals. As a result of the improvement in the measurement of the potential of the discharge, the absolute sizes of the rings agree with the calculated within a mean error of little more than 1 per cent. It is, nevertheless, not certain that the simple theory used in the calculations is exactly true. It has been assumed that the wave-length in the metal of the electrons is the same as *in vacuo*. Bethe has pointed out[*] that the 30 per cent. discrepancy in Davisson and Germers' results could be explained by supposing a refraction of the waves, the index of refraction for the metal being less than unity. Such an effect has been observed for X-rays to a slight extent, and its effect on the observed position of the diffracted beams has been calculated by Darwin.[†] On the de Broglie theory $\mu = \sqrt{(E - V)}/E$, where E is the energy of the electron and V the mean potential in the metal, both in volts. To fit Davisson and Germers' results, V is of the order 20 volts, so that for $E = 40{,}000$, μ differs from unity by 1 in 4000.

---

[*] 'Naturwiss.,' vol. 15, p. 786 (1927).

[†] 'Phil. Mag.,' vol. 27, p. 318 (1913).

In my experiments the electrons are incident normally on the film and so are not deflected at entrance, but the wave-length will change from λ_0 to λ_0/μ. The condition for reflection is $2d \sin \theta = n\lambda_0/\mu$. The electrons are refracted on emergence and the total deviation $2\theta'$ is given by $\sin 2\theta' = \mu \sin 2\theta$. If θ is small, $2\theta' = n\lambda_0/d$ independent of μ; thus the effect would be undetectable even apart from the smallness of $\mu - 1$.

The absence of any appreciable difference in velocity between the electrons of the main and diffracted beams is, perhaps, not surprising. Since the whole crystal influences the diffracted beam, the collision which causes the scattering of the electron must be with the crystal as a whole. As this is of practically infinite mass, there will be no transference of energy, and no change in the velocity of the electron. If, on the contrary, the ray were scattered by another free electron, we should have $u - u' = u \sin^2 \theta$, where u, u' are the initial and final component velocities along the original beam, and θ is the angle through which the *relative* motion has turned. For small angles $\sin \theta = 2 \sin \phi$ when ϕ is the angle of deflection of the ray. Thus $(u - u')/u = 4 \sin^2 \phi$. For the rings considered ϕ is about 0·04, thus $(u - u')/u$ is nearly 1 per cent., and should have been detected. The energy given to the free electron would be $(0{\cdot}04)^2$ of the original energy. This would be of the order 100 volts. Probably a collision with a bound electron would lead to much the same result if this energy were sufficient to ionise, as would be the case for the outer electrons. Thus the experiment serves to confirm the expectation that, if the scattering is regarded as a collision phenomenon, the collisions must be regarded as being with the crystal, or at least the atoms, rather than with electrons. On the other hand, it is known that the loss of energy of electrons passing through matter can be calculated with fair accuracy as due to collisions with the electrons in the matter. It is possible that only a proportion of the scattered electrons are scattered according to the diffraction laws. In practice there is always a great deal of general scattering, and while this may, perhaps, be due to multiple scattering caused by the films being rather too thick, I am inclined to doubt whether this explanation can account for it all. It is hoped that quantitative measurements of the intensity of the rings may throw light on this point.

The question of the number of waves accompanying an electron is of great theoretical importance. The large resolving power seems definitely to show that the waves extend with appreciable amplitude over a distance at least of the order 50×10^{-8}, which makes the wave system of the free electron larger than an atom. The only way of reconciling the experiments with an electron, the amplitude of whose waves is concentrated in a small region, seems to be to suppose that the different electrons in the beam influence each other, so as to produce a coherent wave.* This would mean that a *single* electron, however well known its velocity might be, would not necessarily fall exactly on any of the possible rings. It seems extremely unlikely that the diffraction pattern would become more blurred if the intensity were reduced till the electrons followed each other at such wide intervals that there was only one in the apparatus at once. Even as it is, the chances are greatly against more than one electron being in the film at one time with the current densities I am using. We are led to the conclusion that the appropriate wave, even for a single electron, must be approximately of the form $a \cos 2\pi\nu\,(t - vz/c^2)$. It cannot, however, be exactly of this form, for a "homogeneous" electron, like homogeneous light, would imply a steady motion lasting from minus infinity to plus infinity. The more nearly it approaches this form the more definite is the speed of the electron.

Any arrangement which would, on the older view, have been described as a means of producing electrons of constant velocity can be regarded as a means of producing a long train of electron waves, and there is an exact analogy between electron waves and light waves. The separation of a group of electrons of nearly constant velocity by a magnetic field and a slit is analogous to the production of approximately homogeneous light by a prism spectroscope. Again, since the speed of electron waves increases with the wave-length, there is dispersion even in the absence of a field, and a group of waves grows longer as it advances. This is analogous to the spreading out of a pulse of white light when it goes through a parallel sided slab of glass. To take an analogy from water waves: if a stone is thrown into water the irregular disturbance produced will, when it reaches a great distance from the source, consist of a long train of waves, each differing very slightly from the next, but gradually changing in wave-length. In the same way a group of electrons, which could be represented near the source by a short train of waves, will at a great distance have spread out owing to the difference in speed of the electrons, and those in any small region will have nearly the same velocity, *i.e.*, will correspond to a long train of nearly uniform waves. By such methods as these we can obtain a group of electron waves as nearly homogeneous as we like. Darwin* has shown that in a problem involving the motion of electrons, the electrons need only be considered as particles at the beginning and end where the energy interchanges occur. In between, the problem can be treated purely as a matter of wave theory. For reasons of continuity it seems desirable to retain the conception of electrons at intermediate points. This requires that an electron shall be regarded as associated with a train of waves whose length is determined by the resolving power of the process that has been used to make the electrons homogeneous. With the present apparatus, if the voltage could be regarded as constant, *e.g.*, by the use of a high tension battery, the limit would come from the uncertainty of the velocity with which the electrons are freed from the cathode. Darwin* has shown that a constant acceleration produces no change in the length (in centimetres) of the wave train, though as the wave-length is diminished it increases the number of waves in the train. Still more homogeneous electrons could be obtained (if the diffracting crystals were large enough) by selecting the electrons which were diffracted to a particular point on one of the rings, just as light may be made homogeneous by a diffraction spectroscope used as a monochromatic illuminator. The regular train of waves is produced in this case by the regularities of the crystal.†

---

* [*Note added in proof.*—L. de Broglie ('J. de Physique,' vol. 8, p. 225 (1927)) has shown that an expression of the form $f(x, y, z - vt) \cos 2\pi\nu\,(t - vz/c^2)$ satisfies a modified wave equation, and so represents a permanent group, if $\nabla^2 f = 0$. He takes f as $1/r$ and gets a group with a point singularity. To deal with the motion of such special groups he employs the "principle of double solution," which is equivalent to assuming that the singularities move along the rays of an associated wave of *constant* amplitude. If this is accepted, point electrons are possible, but the difficulties mentioned below are merely transferred to the new wave.]

It should be noticed that in these experiments we must regard the influence of the electron as extending over an area of the crystal equal to the cross-section of the fine tube which defines the electron beam, though, owing to the very small wave-length, a much less area would give a diffraction pattern indistinguishable in practice from that observed.

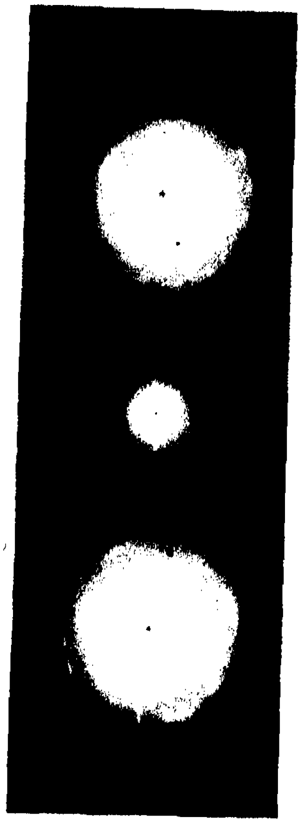
I have again to thank Messrs. C. G. Fraser and J. D. McKay for very valuable assistance in the conduct of these experiments.

---

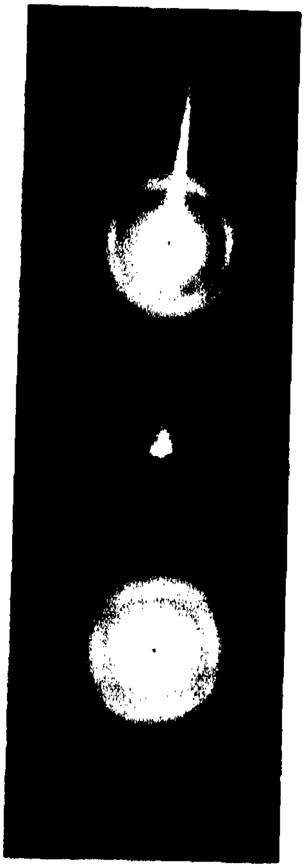
* 'Roy. Soc. Proc.,' A, vol. 117, p. 258 (1928).

† While the number of waves associated with even a single electron can apparently be made indefinitely large, it is a very interesting question as to what can be predicted of a *single* electron with a known short train of waves. If the law connecting velocity and wave-length is exact, such an electron having no definite wave-length can have no definite velocity. Its experimental behaviour is not exactly predictable, an instance of Heissenberg's uncertainty relation, and indeed it seems doubtful how far it can fairly be said to have a velocity at all.

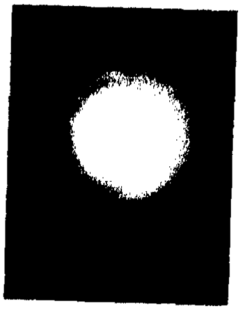
Thomson. *Roy. Soc. Proc., A, vol.* 119, *Pl.* 12A.



(1) Aluminium.

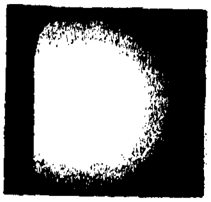


(2) Platinum.

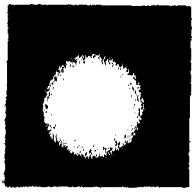


(3) Silver, inclined about horizontal axis.

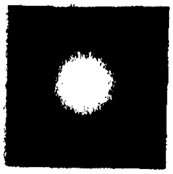
Reid. *Roy. Soc. Proc., A, vol.* 119, *Pl.* 12B.



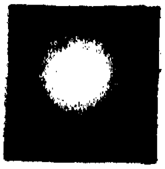
1



2



3



4

(*Facing p.* 602.)

Summary.

(1) Diffraction patterns have been obtained from electrons passing through thin films of platinum which are in complete agreement with those predicted on the de Broglie wave theory.

(2) The values of the grating constants for aluminium, gold and platinum, calculated on the de Broglie theory, agree to 1 per cent. with those found by X-ray analysis.

(3) The velocity of the diffracted electrons differs by less than 1 per cent. from that of the main beam.

(4) The fact that certain rings can be resolved shows that the electrons must be accompanied by a train of not less than about 50 waves in length.

(5) Experiments with inclined films are in qualitative agreement with theory.

(6) The theoretical consequences of some of these results are discussed.

The Diffraction of Cathode Rays by Thin Celluloid Films.

By Alexander Reid, M.A., Assistant in Natural Philosophy at the University of Aberdeen.

(Communicated by Sir Joseph Thomson, F.R.S.—Received May 14, 1928.)

[Plate 12b.]

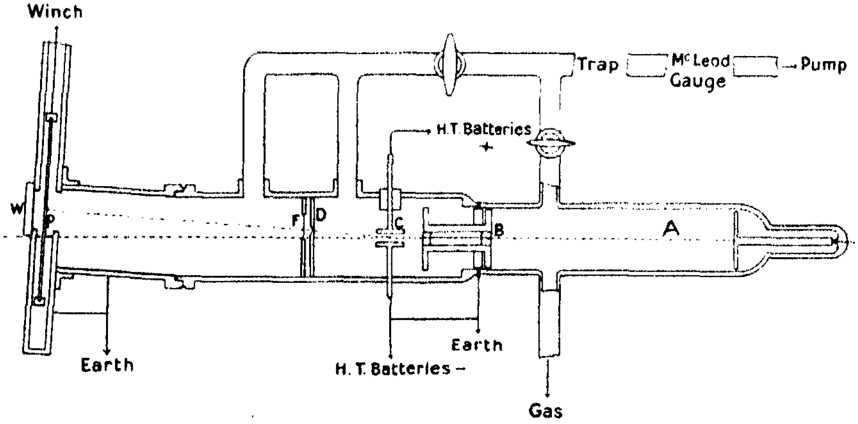
In the de Broglie theory of mechanics a moving particle behaves as a group of waves of wave-length $= h\sqrt{1 - v^2/c^2}/m_0 v$ and phase velocity $V = c^2/v$, where m_0 is the mass of the particle, v its speed and c the velocity of light. In the case of electrons of energy 10,000 to 40,000 volts, the corresponding wave-lengths would be from $1 \cdot 22 \times 10^{-9}$ to $0 \cdot 66 \times 10^{-9}$ cms. These waves, therefore, should behave in many ways like hard X-rays, and, in particular, should show similar diffraction patterns when passed through crystals. Thus, according to the Bragg formula, if the rays are incident at an angle θ on a set of parallel planes, we should get reflection provided that $2d \sin \theta = n\lambda$, d being the spacing between the planes.

To investigate these effects, a series of experiments, suggested and supervised by Prof. G. P. Thomson, was begun in October, 1926. An account of these has already been published in 'Nature,'* but since then much more accurate work has been done. The experiments consisted in sending a beam of homogeneous cathode rays through a thin film at normal incidence and receiving them on a photographic plate. As shown below, the resultant pattern should be of the Hull-Debye-Scherrer type. Celluloid was chosen because it is comparatively easy to get films of the order of 500 Å.U. thick. These were prepared by dissolving celluloid in amyl acetate. A little of the solution was dropped on water and the amyl allowed to evaporate. The celluloid which remained was removed on cardboard frames. It is essential that the films be thin enough to prevent blurring of the pattern by multiple scattering.

---

* 'Nature,' vol. 119, p. 820 (1927).

The apparatus was so designed that the energy of the rays could be accurately determined. Cathode rays, generated in the discharge tube A pass through a fine tube in the anode B and between the plates of a condenser C, where they are deflected upwards by an electric field. Part of the electric spectrum is selected by a small hole in a brass diaphragm D, behind which the film F is mounted. The diameter of the hole is 0·25 mm. and is 2·5 mm. above the straight through path of undeflected rays. The plate P is 20 cms. from the film and is lowered by means of a winch. Three exposures can be taken on one plate. The rays can be examined on a willemite screen W before the plate is lowered. A constant factor connects the energy of the rays passing through the hole in the diaphragm with the voltage applied between the condenser plates, and is easily calculated from the dimensions of the apparatus.



In the earlier experiments considerable trouble was caused by the effect of the earth's field. This deflected the rays sideways past the tube in the anode, and only stray rays were received on the photographic plate. This difficulty was overcome by neutralising the earth's field with Helmholtz coils, and the apparatus then gave very satisfactory results.

Specimens of the photographs are shown on Plate 12B. It will be seen that these all consist of a central spot surrounded by concentric rings, each of which is uniform round the circumference. It is difficult to obtain exposures which give several rings. Nos. 1 and 2 show the third ring with rays of different energies. In No. 3 three rings can be seen, the second of which, perhaps, is less clearly defined than the other two.

If these rings are in the nature of a diffraction pattern, the diameters should be proportional to the wave-length. The latter, from the de Broglie formula, is given by the relation $\lambda = h\sqrt{1 - v^2/c^2}/m_0 v$. v, the velocity of the rays, may be calculated from the equation $\frac{1}{2}\, mv^2 = Pe$, where P is the potential difference in E.S.U. through which the rays have fallen to acquire their energy. Since v/c is small, we have $\lambda = h\sqrt{Pem_0}$. If P is measured in volts, this becomes $\lambda = h\sqrt{150}/Pem_0$. The wave-length is therefore inversely proportional to the fall of potential, and if D is the diameter of any ring, $D\sqrt{P}$ should be constant.

The following table shows how far results are in agreement with this law. D_1, D_2, D_3 are the diameters of successive rings measured in millimetres, L the distance between the plate and film in centimetres.

Plate.	Voltage P.	D_1.	D_2.	D_3.	$D_1\sqrt{P}/L$.	$D_2\sqrt{P}/L$.	$D_3\sqrt{P}/L$.
		mm.	mm.				
March 16 (1)	9,800	12	—	—	54·4	—	—
March 19 (2)	10,800	10·6	—	—	57·5	—	—
March 15 (2)	14,800	9	18·4	—	54·5	112·0	—
March 15 (3)	19,700	8·2	—	—	57·5	—	—
March 15 (1)	26,000	7	14	22·5	56·4	112·9	180·8
March 16 (2)	29,500	6·5	13·6	—	58·8	116·8	—
March 19 (1)	35,100	6·25	12·2	19·8	58	114	186·5
April 17	40,400	5·5	11·5	18·5	55·1	115·8	186·5
				Mean	56·5	114·4	184·6

Prof. Thomson also has taken several photographs with celluloid films. In these the energy of the rays was measured by spark gap. The average value of $D\sqrt{P}/L$ in the same units was 57 as against 56·5 when the energy was measured by electrostatic deflection. This remarkably good agreement shows that spark-gap measurements of voltages are quite accurate, and justifies the assumptions made in employing this method.

Unfortunately, very little is known of the structure of celluloid, but probably it is somewhat similar to the long chain compounds investigated by Müller.* In all of these were found two short spacings independent of the carbon content of the compound and one long spacing. This means that the carbon atoms are arranged in chains, and these are packed with their axes parallel. The mass density is greater along a chain axis than along any other straight line. Of course, the short spacings in celluloid need not be the same as in these fatty acid compounds. In very thin films the cells are probably arranged with their chain axes perpendicular to the surface. This fact is confirmed by Adam in the case of thin films of fatty acids spread on water.* Planes containing sufficient atoms and at the same time giving small enough values of θ to satisfy the relation $n\lambda = 2d \sin \theta$ will be those containing chain axes. The angle θ will be provided by the slight divergence of the beam and lack of evenness of the surface of the film. Prof. Thomson in his paper on the scattering of cathode rays by thin metal films† showed that if the undeflected rays strike the plate at O, rays reflected by a plane making an angle θ with the primary beam strike the plate at a point P, where $OP = 2\theta L$, L being the distance of the plate from the film. Since planes perpendicular to the surface of the film can be taken all round the incident beam, we get a ring on the plate of diameter $= 2OP = 4\theta L$. Below are tabulated a few of the results :—

---

* 'Roy. Soc. Proc.,' A, vol. 114 (1927).

Plate.	D_1.	D_2.	D_3.	D_2/D_1.	D_3/D_1.
	mm.	mm.	mm.		
May 3 (2)	4·9	9·6	15·5	1·96	3·16
	3·9	7·4	12	1·9	3·09
November 4	7·2	14·1	22	1·96	3·06
December 3	6·8	13·9	—	2·02	—
March 19	6·25	12·2	19·8	1·95	3·17
April 17	5·5	11·5	18·5	2·01	3·36
			Mean	1·96	3·17

From this table it will be seen that the rings are in the ratio 1 : 2 : 3. No plate was obtained where this was not the case. We are justified, then, in assuming that these rings are first, second and third order spectra from the same set of parallel planes.

From the foregoing data, the spacing of the planes giving the pattern may be calculated. The mean value for several results was 4·46 Å.U. The rings with celluloid were never so sharp as those obtained by Prof. Thomson with metal films. A possible explanation of this is that celluloid, like the fatty acids, may have two nearly equal spacings. The blackening on several of the plates was measured photometrically and the percentage of rays scattered

* 'Roy. Soc. Proc.,' A, vol. 99, p. 336 (1921), and vol. 101, p. 452 (1922).

† 'Roy. Soc. Proc.,' A, vol. 117, p. 600 (1928).

plotted against position along a line through the central spot. In the best of these curves, two maxima about 0·8 mm. apart were found corresponding to what had been taken as a single ring. Assuming that these maxima mark the positions of two nearly equal rings, the corresponding spacings are 3·67 Å.U. and 4·35 Å.U. The values for fatty acids given by Müller are 3·67 and 4·08 Å.U.

It is fairly conclusive, I think, that these photographs must be accepted as diffraction patterns, and the close agreement of experimental results with those predicted by the de Broglie theory offers strong evidence in its favour.

I should like to express my thanks to Prof. Thomson, through whose interest in and supervision of the work much of the success of these experiments is due, and to Mr. C. G. Fraser and Mr. J. D. McKay for their assistance in handling the apparatus.

Summary.

1. Experiments are described in which diffraction patterns were obtained by passing cathode rays through thin celluloid films.

2. The energy of the ray was measured accurately by electrostatic deflection, and the results obtained agree to within 1 per cent. with those of Prof. Thomson, who used a spark gap to measure the voltage.

3. A probable structure of the films is suggested, and the close agreement of experimental results with those predicted by the de Broglie theory of wave mechanics is indicated.

The Diffraction of Cathode Rays by Thin Films of Copper, Silver and Tin.

By R. IRONSIDE, M.A., Robbie Scholar of Aberdeen University.

(Communicated by Sir Joseph Thomson, F.R.S.—Received May 14, 1928.)

The experiments to be described are a continuation of the work done by Prof. G. P. Thomson on the scattering of cathode rays by thin films. The details of the apparatus will be found in his paper.*

Procedure.

Cathode rays are generated in hydrogen in a discharge tube by an induction coil and pass through a fine tube at the end of which a film is mounted. The rays pass through this film and travel along the camera to the photographic plate which, when an exposure is to be made, is lowered into the path of the rays. The camera is exhausted along with the tube to a low vacuum by a three-stage mercury diffusion pump. The camera and anode are earthed, and the spark gap between two 4 cm. brass balls, which are connected directly across the tube, is read to give the potential through which the electron falls. A mercury break acts as interrupter in the primary and a rectifying valve is used in the secondary.

Introduction.

According to the de Broglie theory a moving particle behaves as a group of waves, the wave-length being given by $\lambda = h\sqrt{1-\beta^2}/m_0 v$, where m_0 is the mass for slow speed, v is the velocity of the moving particle, $\beta = v/c$, where c is the velocity of light and h is Planck's constant. The waves in passing through a crystal will give a Laue pattern just as X-rays do, but as they are far less penetrating than X-rays very thin films must be used. If the film consists of a large number of minute crystals orientated at random, the diffraction pattern will show a series of rings like those got in X-ray powder analysis. If, however, it consists of a few larger crystals set in a definite direction, the circles will be modified by contracting to diametrically opposite arcs resembling more the original Laue patterns. Both these types of patterns have been obtained.

By the Bragg formula we have $n\lambda = 2d \sin\theta$. If D be the diameter of the ring in the diffraction pattern and L the length from film to plate $\sin\theta = \mathrm{D}/4\mathrm{L}$

---

* 'Roy. Soc. Proc.,' A, vol. 117, p. 600 (1928).

approximately, the average error made by assuming the relation to be exact being about 0·04 per cent. If P is the potential difference in volts through which the rays fall to acquire their energy $\lambda = h\sqrt{150/ePm_0}/(1 + eP/1200m_0c^2)$. Since λ is proportional to D, $D\sqrt{P}(1 + Pe/1200\,m_0c^2)$ should be constant. Good evidence for this has been found in these experiments upon silver, copper and tin. Moreover, in working with substances whose crystal constants are known, we can verify the hypothesis by comparing the values found by experiment with those given in X-ray analysis.

Silver.

Silver films sufficiently thin were not difficult to obtain. One was deposited chemically in the process of silvering a mirror; others were prepared by reducing silver leaf in dilute nitric acid. Some values of

$$D_1\sqrt{P}(1 + Pe/1200\,m_0c^2)$$

are shown below :—

Plate.	S.G.	P.	D_1.	$D_1\sqrt{P}(1 + Pe/1200\,m_0c^2)$.
	mms.	volts.	cms.	
December 23 (1)	14·5	43,600	1·50	320
December 23 (2)	17	50,000	1·40	322
December 23 (3)	18·5	53,000	1·40	328
December 23 (4)	21	56,900	1·30	318
				Mean 322

$D_2\sqrt{P}$ and $D_3\sqrt{P}$ are also fairly constant. The following table gives the actual sizes of the rings and the ratios of the diameters of the other rings to that of the first. Part of a fifth and a sixth ring appear, but the plate is too small to show the whole of these.

Plate.	D_1	D_2	D_3	D_4	D_2/D_1	D_3/D_1	D_4/D_1
	cms.	cms.	cms.	cms.			
December 23 (1)	1·50	1·76	2·47	3·02	1·17	1·65	2·01
December 23 (2)	1·40	1·62	2·34	2·86	1·16	1·67	2·04
December 23 (3)	1·40	1·61	2·31	2·80	1·15	1·65	2·00
December 23 (4)	1·30	1·51	2·19	2·63	1·16	1·68	2·02
Mean					1·16	1·66	2·02

Silver is a face-centred cubic lattice, therefore the diameters of the rings should be as $\sqrt{3} : \sqrt{4} : \sqrt{8} : \sqrt{11} : \sqrt{12}$. The $\sqrt{11}$ and $\sqrt{12}$ rings are too close to be separated, so the ratios of the diameters of the other rings to D_1

should be 1·15, 1·63 and 1·96 with which the values obtained are in good agreement. None of these rings are absent and the unbroken nature of the rings shows that the films consist of a conglomeration of minute crystals with no definite orientation.

The values of a, the lattice-constant, found from both kinds of films are in good agreement. The average value found is $a = 4 \cdot 11 \times 10^{-8}$ cm., while X-ray analysis gives $4 \cdot 079 \times 10^{-8}$ cm., showing a difference of less than 1 per cent.

Copper.

Thin copper films were more difficult to prepare. The thinnest copper leaf obtainable was etched down by floating in nitric acid. Some films first made gave only two vague, broad rings with traces of a third. The average ratio of the diameters of the first and second rings is 1·68, which is almost $\sqrt{8}/\sqrt{3}$. With films thin enough to be transparent there were seven measurable rings. The table below gives the values of $D \, . \chi$ for the various rings where χ is written for $\sqrt{P}\,(1 + Pe/1200\, m_0c^2)$.

Plate.	S.G.	P.	$D_1 \cdot \chi$	$D_2 \cdot \chi$	$D_3 \cdot \chi$	$D_4 \cdot \chi$	$D_5 \cdot \chi$	$D_6 \cdot \chi$	$D_7 \cdot \chi$
	mms.	volts							
December 20 (1)	17·5	51,100	—	—	329	426	514	610	—
December 21 (1)	15	45,000	200	—	312	426	509	596	—
December 21 (2)	12·5	38,500	188	—	309	423	502	594	—
December 21 (3)	10	31,800	194	—	309	416	500	—	—
December 21 (4)	7·25	23,800	—	—	311	427	493	—	—
December 21 (6)	19	54,200	—	—	331	432	522	611	735
December 22 (1)	22·5	62,400	—	—	334	435	526	623	732
December 22 (4)	17·5	51,100	—	248	312	431	517	614	748
December 22 (5)	20	56,500	—	253	314	427	516	620	—
December 22 (6)	22	61,000	—	264	325	437	531	629	726
Mean			194	255	318	428	513	612	735

Plate.	D_3/D_2	D_5/D_3	D_6/D_4	D_7/D_4
December 20 (1)	—	1·57	1·43	—
December 21 (1)	—	1·63	1·40	—
December 21 (2)	—	1·62	1·41	—
December 21 (3)	—	1·61	—	—
December 21 (4)	—	1·58	—	—
December 21 (6)	—	1·59	1·42	1·68
December 22 (1)	—	1·54	1·42	1·67
December 22 (4)	1·24	1·64	1·43	1·75
December 22 (5)	1·24	1·65	1·45	—
December 22 (6)	1·23	1·64	1·45	1·66
Mean	1·24	1·61	1·43	1·69

The ratios between D_1 and the diameters of the other rings are not those between the square roots of any small integers. Among the other rings we have the relations shown by the table of ratios. Thus there are two sets of rings: the second, third and fifth; and the fourth, sixth and seventh; and there is no simple connection between the two sets. Now, $D_6/D_4 = 1 \cdot 43 = \sqrt{2}$ and D_4 is just the right size of ring for diffraction due to the (2, 0, 0) planes of a copper crystal. So D_6 gives a (2, 2, 0) plane. We expect two unresolved rings from the planes the sum of the squares of whose indices are 11 and 12, but as the latter is the second order of reflection from a (1, 1, 1) plane, and the first order is absent, it is rejected. Professor Thomson has explained the absence of the ring due to a (1, 1, 1) plane in aluminium by the orientation of the crystals in the process of rolling the film so that one axis is perpendicular to the film. If the ray is parallel to the x axis, a (1, 1, 1) plane will make too big an angle to give interference, while a (3, 1, 1) plane makes an angle of 13°. The (2, 0, 0) and (2, 2, 0) planes give rings four and six, and as we would expect both rings consist of four circular arcs at the ends of two perpendicular diameters, the diameters of the one set being inclined at 45° to the others. As these arcs are fairly long there is considerable variation in the directions of the crystals which may be great enough to enable a (3, 1, 1) plane to produce its ring. This seems certain when we compare D_7/D_4, which is $1 \cdot 69$, with $\sqrt{11}/2 = 1 \cdot 66$. This ring, as would be expected, is broken into several arcs.

In the other set of rings $D_3/D_2 = 1 \cdot 24$ and $D_5/D_3 = 1 \cdot 61$. These ratios are very near the values of $\sqrt{3}/\sqrt{2} = 1 \cdot 22$ and $\sqrt{8}/\sqrt{3} = 1 \cdot 63$, which again suggest a cubic lattice. D_3 and D_5 may be due to the copper atoms in cuprous oxide which are arranged on a *face*-centred lattice, but D_2 would have to be due to a (1, 1, 0) spacing, which is excluded. The oxygen atoms, however, form a *cube*-centred lattice of the same lattice-constant for which the $\sqrt{2}$ ring is the most prominent. The third and fifth rings, then, are due to (1, 1, 1) and (2, 2, 0) planes; the third must be accounted for by the copper atoms alone, while the fifth may be caused in part by the oxygen atoms as well. The (2, 0, 0) planes of both ought to give a ring intermediate to the third and fourth, but only in the last plates in which the second ring appeared strongly was there a trace of this ring. The oxygen atoms ought also to give a ring corresponding to a (2, 1, 1) spacing, but this ring would certainly be weak and too near the fourth to be distinguished from it. The second, third and fifth rings, like those of silver, are quite continuous and unbroken. They are more pronounced than the rings due to copper and grow progressively stronger with the increasing age of the film and consequent oxidation.

There are traces of other two outer rings on some plates.

The first ring, which soon disappears, is no doubt due to some organic substance as the spacing is the same as that of celluloid and another substance (probably organic) investigated by Prof. Thomson.

The data from D_4 and D_6 (D_7 is not so well marked) give an average lattice-spacing for copper of $3 \cdot 66 \times 10^{-8}$ cm., while D_3 and D_5 (D_2 is faint and appears with one film only) give $4 \cdot 30 \times 10^{-8}$ cm. for cuprous oxide. The values given by X-rays are $3 \cdot 60 \times 10^{-8}$ cm. and $4 \cdot 26 \times 10^{-8}$ cm., showing a discrepancy of 1 or 2 per cent.

Tin.

Not much success has been obtained with tin owing to the difficulty in making films. Tin-foil was beaten out between gold-beater's skin with polished hammer and anvil and then allowed to thin down in acid—acetic acid giving the best results. The films were lifted, washed free of acid and mounted, but the thinnest were always broken in drying. None were obtained thin enough to give more than one ring with sometimes a broad second ring. As before $D\sqrt{P}\,(1 + Pe/1200\, m_0c^2)$ is fairly constant, the value being somewhat above the average when P is small.

Plate.	S.G.	P.	D.	$D\sqrt{P}\,(1 + Pe/1200\, m_0c^2)$.
	mms.		cms.	
March 6 (2)	18·5	53,000	1·14	267
March 6 (3)	13·75	41,600	1·32	274
March 13 (2)	14·5	43,300	1·29	274
March 14 (1)	20	56,500	1·12	272
March 14 (2)	21·5	60,000	1·09	274
March 14 (3)	9	28,900	1·69	289
				Mean 275

White tin is of tetragonal structure, the spacings d_{200} and d_{020} being $2 \cdot 91 \times 10^{-8}$ cm., while the spacing calculated from the mean value, 275, above is $2 \cdot 86 \times 10^{-8}$ cm. The average ratio of the second ring, when measurable, to the first is 2·0, so it is the second order of the first ring.

I have to thank sincerely Prof. G. P. Thomson for guidance in this work, and Mr. C. G. Fraser and Mr. J. D. McKay for much valued assistance in experimental difficulties.

Summary.

These experiments corroborate previous work done on thin films giving support to the de Broglie theory. Silver, copper and tin have been tested with voltages ranging from 29,000 to 62,000. In spite of a large probable error in measuring voltage by means of a spark gap, the quantitative agreement of the results with those of X-rays is good—the discrepancy amounting to only 1 per cent.

The Scattering of Cathode Rays.

By B. F. J. Schonland, M.A., Ph.D., Senior Lecturer in Physics,
The University of Cape Town.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received May 17, 1928.)

§ 1. *Introduction.*

A previous paper by the writer[*] described some measurements of the scattering of cathode rays by thin metal foils. The observations were made under single scattering conditions, and it was concluded that they were in good quantitative agreement with the theory of single scattering due to Rutherford, modified to take account of the alteration in the mass of the particle with its velocity. Such a modification of the orbit of an electron in the field of an atomic nucleus was first suggested by C. G. Darwin,[†] and it was used in this case to derive a "relativity" correction to the original α-ray scattering formula.[‡]

The purpose of this paper is to show that the form of the correction used was not really applicable to these particular experiments, to derive a more complete correction, and to make a fresh comparison of the measurements with theory, though it will be shown that this comparison can only be approximate.

Let p be the length of the perpendicular drawn from the nucleus upon the original path of the particle whose charge is e, and whose mass at low velocity is m_0. Let Ne be the charge on the atomic nucleus and β the ratio of the velocity of the particle to that of light. Darwin has shown that the orbits can then be divided into two classes according as p is greater or less than a critical value p_0 given by

---

[*] 'Roy. Soc. Proc.,' A, vol. 113, p. 87 (1926).

[†] 'Phil. Mag.' [6], vol. 25, p. 201 (1913).

[‡] Crowther and Schonland, 'Roy. Soc. Proc.,' A, vol. 100, p. 526 (1922).

$$p_0 = (\mathrm{N}e^2\sqrt{1-\beta^2})/m_0c^2\beta.$$

Class I orbits occur when p is greater than p_0, and take the form of a modified hyperbola, each element of vectorial angle being increased in the ratio of I to $\sqrt{1-p_0^2/p^2}$. This ratio becomes large if p is near to p_0, and the electron then winds round the nucleus, sometimes making several turns or loops before it can escape.

Class II orbits occur when p is equal to or less than p_0, and take the form of a spiral, the particle winding round and round the nucleus until it falls into it.

These modifications were first applied to the calculation of the scattering of electrons by Crowther and the writer.* In our β-ray experiments, less than 1 per cent. of the scattered particles could have come within the critical distance p_0 of an atomic nucleus, so that it was not necessary to consider the spiral orbits, or the looped orbits of Class I, in deriving the relativity correction. In the cathode ray experiments referred to, however, a neglect of these orbits leads to serious error. The measurements with aluminium foils, for example, were made under conditions such that all the scattered particles approached a nucleus within a distance p less than $2 \cdot 9\ p_0$. The original form of the correction is clearly inapplicable to these experiments, and a more complete form must be used. This will be given later, but it is necessary to discuss two difficulties which arise.

The first is the question of the loss of energy by radiation from the orbits, which may be considerable for such small values of p. It is unfortunate that a direct calculation of this effect seems impossible. A rough and indirect estimate was made by Darwin† on the basis of classical electromagnetic theory, but in view of the success of the quantum theory in this field it cannot be used with any confidence. From the work of Kramers‡ and of Wentzel§ on the quantum theory of " white " X-radiation, it would seem that radiation does occur from orbits of this kind. If so, the orbits of Class I would be modified still further and the amount of scattering increased. At the same time the critical distance p_0 and the number of particles in spiral orbits would increase. The effect of radiation of energy in modifying the orbits and increasing the amount of scattering would be more important in experiments like those of the writer, where the majority of the scattered particles had values of p close to p_0, than in other β-ray experiments.

---

* *Loc. cit.*

† *Loc. cit.*

‡ 'Phil. Mag.,' vol. 46, p. 836 (1923).

§ 'Z. f. Physik,' vol. 27, p. 257 (1924).

In view of the difficulty of forming even a rough estimate of the effect of radiation loss upon the scattering, all that can be done is to calculate the relativity correction which would apply in the absence of radiation, and to see how far this will account for the observed results, bearing in mind that the occurrence of radiation would have the effect of still further increasing the scattering.

On proceeding to this calculation the spiral orbits present a difficulty, for the subsequent history of a particle which spirals inwards into a nucleus is unknown. It was pointed out by Darwin[*] that these orbits must lead to the ultimate escape of the electron from the immediate neighbourhood of the nucleus, for otherwise a transmutation of the elements under cathode ray bombardment would be easily observed. He suggested that the process of escape did not endow the electron with its original velocity, and that this loss of energy might be the main cause of the absorption of the β-rays by matter. More recent evidence does not support this view, and since it is more probable that these spiral orbit particles escape from the atom without serious loss of energy, we may suppose that the direction in which they emerge is independent of their original direction of motion, so that they are scattered equally in all directions. With this assumption as to the spiral orbits and an unavoidable neglect of the possibly important radiation factor, an estimate of the amount of scattering will now be made.

§ 2. *The Complete Relativity Correction neglecting Radiation.*[†]

The equation to the orbit given by Darwin can be written

$$p\gamma/r = \sin\gamma\theta + (1-\cos\gamma\theta)\,p_0/(p\gamma\beta),$$

where

$$\gamma = \sqrt{1-p_0^2/p^2}.$$

If we write $p = p_0 \operatorname{cosec}\psi$, this becomes

$$p\gamma/r = \sin(\theta\cos\psi) + [1-\cos(\theta\cos\psi)]\tan\psi/\beta.$$

If the particle is deflected through an angle ϕ, then $\pi+\phi$ is the angle between the asymptotes and $\frac{1}{2}(\pi+\phi)$ is the angle of minimum r. Hence

$$\beta\cot\psi = \tan[\pi - \tfrac{1}{2}\cos\psi\,(\pi+\phi)],$$

an equation from which ψ can be found if ϕ and β are known.

---

[*] *Loc. cit.*

[†] The present extension of the theory was indicated by Prof. Darwin in the draft of the calculation which he very kindly worked out for me in 1921.

If p_ϕ is the value of p for which the angle of deflection is ϕ, then $p_\phi = p_0 \operatorname{cosec} \psi_\phi$.

If a number Q of the particles fall on a foil of thickness t with n atoms per cubic centimetre, the number of these which have a collision between p and $p + dp$ is $Qnt \times 2\pi p\, dp$ or $Qnt\pi\, (dp^2/d\phi)\, d\phi$, since p is known in terms of ϕ.

Hence the total number of particles scattered through an angle greater than ϕ is $Qnt\pi p_\phi^2$, or $Qnt\pi p_0^2 \operatorname{cosec}^2 \psi_\phi$.

What is usually observed in measurements of the amount of scattering is not, however, the fraction of the main beam which is scattered through an angle greater than ϕ, but the fraction which fails to emerge through a cone of half-angular aperture ϕ. This only includes those rays which have been deflected through angles lying between ϕ and $2\pi - \phi$, $2\pi + \phi$ and $4\pi - \phi$, $4\pi + \phi$ and $6\pi - \phi$, and so on. For example, those rays which are deflected through angles lying between $2\pi - \phi$ and $2\pi + \phi$ are able to emerge with the main beam and are not measured amongst the scattered rays.

The number of particles actually turned into the cone of half-angle $\pi - \phi$ and lost from the emergent beam is thus

$$Qnt\pi\, [p_\phi^2 - p^2_{2\pi-\phi} + p^2_{2\pi+\phi} - p^2_{4\pi-\phi} + \ldots + p_0^2\, 2\pi\, (1 + \cos\phi)/4\pi],$$

or

$$Qnt\pi p_0^2\, [\operatorname{cosec}^2 \psi_\phi - \operatorname{cosec}^2 \psi_{2\pi-\phi} + \operatorname{cosec}^2 \psi_{2\pi+\phi} - \operatorname{cosec}^2 \psi_{4\pi-\phi} \ldots + (1 + \cos\phi)/2].$$

The last term in these expressions refers to those particles which were caught in spiral orbits, and are thrown back, after escape, into the cone of solid angle $2\pi\, (1 + \cos\phi)$.

Substituting the value of p_0 and dividing by Q, the fraction of the beam which will be scattered back is

$$\pi nt\, [Ne^2 \sqrt{1 - \beta^2}/m_0 c^2 \beta]^2\, \Sigma \operatorname{cosec}^2 \psi.$$

The value of this fraction, if relativity orbital modifications are neglected, is

$$\pi nt\, [Ne^2/\{2m_0c^2\, (1/\sqrt{1 - \beta^2} - 1)\}]^2 \cot^2 (\phi/2)$$

and the ratio of the " relativity " to the " non-relativity " expression is

$$[2\, (1 - \sqrt{1 - \beta^2})/\beta \cot (\phi/2)]^2\, \Sigma \operatorname{cosec}^2 \psi = f^2\, (\phi, \beta) \ldots . \qquad (1)$$

This ratio is the relativity correction to the original Rutherford scattering formula. In the form originally given by Crowther and the writer,* only the first term in $\Sigma \operatorname{cosec}^2 \psi$ was considered, since the succeeding terms, which refer to looped and spiral orbits, were not of any importance in the experiments to which the correction was applied.

---

* *Loc. cit.*

§ 3. *Application of the Correction to Cathode Ray Scattering.*

The correction has now to be applied to the cathode ray scattering experiments of the writer, in which case, as has been explained, it is necessary to use the full expression for Σ cosec^2 ψ. The most extensive measurements were made with cathode rays for which β was 0·4967, and the scattering was measured between the angles 90° and 161° 20′. The simplest method of applying the relativity correction is to determine it in the first place for each of these angles separately. Taking first $\phi = 90°$, the values of ψ and cosec^2 ψ have been calculated for ϕ, $2\pi - \phi$, $2\pi + \phi$, etc., and are given in Table I.

Table I.

——	ψ	cosec^2 ψ
$\phi = 90°$	20° 16′	+ 8·334
$2\pi - \phi = 270°$	46° 32′	− 1·899
$2\pi + \phi = 450°$	58° 47′	+ 1·367
$4\pi - \phi = 630°$	65° 36′	− 1·205
$4\pi + \phi = 810°$	69° 57′	+ 1·134
$6\pi - \phi = 990°$	72° 59′	− 1·095
$6\pi + \phi = 1170°$	75° 12′	+ 1·069
$8\pi - \phi = 1350°$	76° 53′	− 1·055
$(1 + \cos\phi)/2$	—	+ 0·500

In the last column each term required for forming Σ cosec^2 ψ has been given the appropriate sign, hence

$$\Sigma \operatorname{cosec}^2 \psi = 6{\cdot}435 + 0{\cdot}162 + 0{\cdot}039 + 0{\cdot}014 \ldots + 0{\cdot}500 = 7{\cdot}15.$$

The first term in this expression represents the number of particles scattered between 90° and 270°, the second the number between 450° and 630°, and so on, while the last term represents the number which emerge between 90° and 270° after spiralling into the nucleus.

Substituting the above value of Σ cosec^2 ψ in the expression for $f^2(\phi, \beta)$ found in the preceding section, the value of the relativity correction for $\phi = 90°$ is 2·022.

The next table gives the corresponding figures for 161° 20′.

Table II.

—	ψ	$\operatorname{cosec}^2 \psi$	
$\phi = 161° \, 20'$	$33° \, 12'$	$+ \, 3 \cdot 334$	
$2\pi - \phi = 198° \, 40'$	$38° \, 32'$		$- \, 2 \cdot 576$
$2\pi + \phi = 521° \, 20'$	$61° \, 56'$	$+ \, 1 \cdot 284$	
$4\pi - \phi = 558° \, 40'$	$63° \, 18'$		$- \, 1 \cdot 253$
$4\pi + \phi = 881° \, 20'$	$71° \, 15'$	$+ \, 1 \cdot 116$	
$6\pi - \phi = 918° \, 40'$	$71° \, 53'$		$- \, 1 \cdot 107$
$(1 + \cos \phi)/2$	—	$+ \, 0 \cdot 026$	

Here

$$\Sigma \operatorname{cosec}^2 \psi = 0 \cdot 758 + 0 \cdot 031 + 0 \cdot 009 \ldots + 0 \cdot 026 = 0 \cdot 824$$

and

$$f^2 (\phi, \beta) = 8 \cdot 62.$$

The fraction of the incident rays scattered outside a cone of half-angle ϕ is

$$\pi/4 \times nt \, (Ne^2/T)^2 \cot^2 \phi/2 \times f^2 (\phi, \beta),$$

where t is the thickness of the foil of atomic number N, and T is the energy of the rays.* The fraction found to be scattered between the angles ϕ_1 and ϕ_2 is therefore

$$\rho = \pi/4 \times nt \, (Ne^2/T)^2 \, [\cot^2 \phi_1/2 \times f^2 (\phi_1, \beta) - \cot^2 \phi_2/2 \times f^2 (\phi_2, \beta)].$$

For comparison with the experimental results for aluminium take $t = 1 \times 10^{-4}$ cm., $N = 13$, $T = 77 \cdot 3 \times 10^3$ electron-volts. The factor in square brackets is $[1 \times 2 \cdot 02 - 0 \cdot 027 \times 8 \cdot 62]$ or $1 \cdot 79$, which gives $\rho = 0 \cdot 00494$. If account is taken of the fact that the portion of the beam which was scattered between 161° 20′ and 198° 40′ fell upon an aluminium cap, which would throw back about 20 per cent., the calculated value of ρ becomes $0 \cdot 00508$.

The value of ρ observed under these conditions† was $0 \cdot 00612$, or $1 \cdot 20$ times that calculated with a neglect of radiation of energy, and on the assumption that the spiral orbit particles are scattered equally in all directions.

If we assume, on the other hand, that the spiral orbit particles are captured by the atom, the last term in the expansion of $\Sigma \operatorname{cosec}^2 \psi$ must be suppressed. The calculated value for the amount of scattering then becomes $0 \cdot 00474$, making the ratio of the observed to the calculated value $1 \cdot 29$.

The measurements which were made to test the variation of the amount of scattering with the thickness and atomic number of the foil, as well as with the energy of the rays, require re-examination in the light of this new form for $f^2 (\phi, \beta)$, and the attendant uncertainty as to the effect of radiation.

* Crowther and Schonland, *loc. cit.*

† Schonland, *loc. cit.*

It appears that the first two tests, in which it was found that ρ/t and ρ/nN^2 remained constant, are still valid, for $f^2(\phi, \beta)$ is independent of t and N, and should remain so even when modified to take account of radiation. On the other hand, the variation of ρ with the energy of the rays involves $f^2(\phi, \beta)$, and no satisfactory test can be made if the radiation factor is unknown. It may be noted, however, that the constancy of the quantity ρ^{T^2}/α, where $\alpha = [\cot^2 \phi_1/2 \times f^2(\phi_1, \beta) - \cot^2 \phi_2/2 \times f^2(\phi_2, \beta)]$, which was found to hold fairly well in the experiments, is improved by using the new form for $f^2(\phi, \beta)$. The old and the new values of α are compared in the last column of the following table, which also contains relativity corrections calculated for three different velocities and two different angles.

Table IV.

β	$f^2(\phi, \beta)$		a (new).	a (old).	a (new)/a (old).
	$\phi = 90°$	$\phi = 161° 20'$			
0·497	2·02	8·62	1·79	2·26	0·796
0·400	1·66	6·71	1·48	1·83	0·809
0·333	1·47	5·64	1·32	1·56	0·846

§ 4. *Summary and Discussion.*

It has been shown that the conditions prevailing in the writer's cathode ray scattering experiments are unsuitable for an adequate test of the relativity modification to the original scattering formula, mainly because a further modification, the effect of which cannot be calculated, is probably caused by radiation from the scattered electrons. The full relativity correction, disregarding radiation loss, has been developed and used to obtain an estimate of the amount of scattering to be expected. This is an under-estimate, for if radiation takes place the scattering will be increased. The ratio of the observed to the calculated amount of scattering is found to be 1·20 on a reasonable assumption as to the subsequent history of an electron describing a spiral orbit.

Without over-emphasising the value of this estimate, it may be pointed out that the original Rutherford α-ray theory of scattering, without the modifications required by the change of mass with velocity or by a possible radiation of energy, gives a value for the amount of scattering which is only 0·44 of that observed.

Any experiment to test the modified theory of the scattering of β or cathode rays in an adequate manner must be carried out under conditions such that the perpendicular distance p from the nucleus upon the original path of the particle is large compared with the critical distance p_0. With the voltages so far available in cathode ray experiments, and the thinnest foils obtainable, this is only possible if the angle of scattering is small, when, however, the scattering observed is no longer single. To steer between the Scylla of spiral orbits and radiation loss and the Charybdis of plural scattering requires the use of fast β rays and moderate angles of scattering, such as were employed in the β-ray experiments of Chadwick and Mercier.*

§ 5. *Summary.*

It is shown that the writer's recent cathode ray scattering experiments were not suitable for an adequate test of the relativity correction to the orbit of a β-particle deflected by an atomic nucleus, owing to the effect of radiation and of the spiral form the orbit may take. A full treatment of the effect of the relativity orbits is given, and an estimate is made of the amount of scattering to be expected which is in fair agreement with that observed.

Metal Overvoltage Measurements with the Cathode Ray Oscillograph.

By Edgar Newbery, D.Sc., F.I.C., University of Cape Town.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 24, 1928.)

[Plate 13.]

Previous work† by the author on metal overvoltages, carried out with the aid of a rotating commutator, led to the conclusion that with certain well-marked exceptions, metal overvoltages, both cathodic and anodic, were in the neighbourhood of zero.

In view of the objections raised by certain workers to results obtained with the rotating commutator, it was decided to check these results with the cathode ray oscillograph, and the following work was therefore carried out with this object in view. It was also done in the hope of providing further data for the solution of the problem of the origin of transfer resistance.

* 'Phil. Mag.,' vol. 50, p. 208 (1925).

† 'J. Chem. Soc.,' vol. 111, p. 470 (1917).

Experimental.

The apparatus used was identical with that previously described, with one important exception to be described later. The electrodes used were copper, silver, zinc, cadmium, mercury, lead, iron, nickel and cobalt, both anode and cathode being of the same metal. The electrolytes were normal solutions of the corresponding metallic sulphates, except in the case of silver, lead and mercury, where the nitrates were used. In the case of iron, ferrous and ferric sulphates were used.

Preliminary experiments with the first four metals confirmed previous observations in so far as metal overvoltages were concerned, these being nearly zero in each case; but at the same time these experiments showed a very appreciable transfer resistance in marked disagreement with previous experiments by another method.* After a somewhat lengthy search for the reason for this disagreement, suspicion became concentrated upon the design of the electrode vessel. This was a rectangular glass cell, 6 cms. long, 1 cm. wide and 10 cms. deep. In this the anode, a strip of metal 1 cm. wide and 5 cms. long, occupied one end, and the cathode, 1 cm. wide and 1 cm. long, was fixed about 1 cm. away, both electrodes fitting the vessel from back to front fairly closely. It was imagined that the lines of flow of the current would be confined almost entirely to one end of this cell, especially as the back of the cathode was insulated with sealing-wax, and this appeared to be confirmed by the fact that the point of the standard electrode vessel could be moved about anywhere behind the cathode without any appreciable effect on the curves.

The first indication that the above supposition was incorrect was given when it was observed that if the distance between anode and cathode was reduced to 0·5 cm., the apparent transfer resistance was reduced slightly but distinctly. Again, it was found that this apparent transfer resistance was nearly constant with varying current density, and, further, that in solutions of different conductivity it varied directly with the resistance of the solutions.

In order to avoid the interference of this strong field of force in the electrolyte, two methods were tried, both of which appear to fulfil this object.

In the first place, a cell was constructed as shown in fig. 1. The tubular of A projecting into C had a cross section of 1 square cm., and the end was carefully ground flat. The experimental electrodes were made in the form of small discs of the same size as this end of the tube, and were pressed up against it during the experiment. The two vessels were cemented together with hard white sealing-wax, and the back of the cathode covered with similar material. When the anode is placed in A and the extremity of the standard electrode vessel in C, it is evident that any appreciable leakage of current between the two is effectively prevented.

---

* 'Trans. Far. Soc.,' vol. 15, p. 126 (1919).

The second method consisted in drawing out the end of the standard electrode vessel to a fine jet bent upwards in a U, and then slightly outwards at the top. The tip of this jet was then ground flat in a vertical plane, and pressed against the surface of the experimental electrode. This type of electrode tip has been used by many other workers in attempts to measure overvoltage by the direct method.

Since metal overvoltages must be measured from the single potential of the metal itself as standard, this single potential was marked on the photographic plates by passing a small constant current through the deflecting solenoids from a subsidiary battery, and exposing the plate when no current was passing through the experimental cell. This was done before and after each experiment, but in most cases the second spot thus obtained was exactly superimposed upon the first. For convenience these spots will be referred to as S.P. 1 and S.P. 2, respectively. They are seen between the voltage marking spots and the main curves on the diagrams.

Results.

Fig. 2 is a curve obtained with copper in a normal solution of copper sulphate, and is typical of all the curves obtained with metals not in the iron group. The current density was 10 milliamperes per square centimetre intermittent (≐ 20 m.a. continuous), and it is at once seen that both overvoltage and transfer resistance are nearly zero. The slight irregularity in the line is fully accounted for by the presence of traces of hydrogen liberated along with the copper under the conditions of the experiment. Exactly similar curves are given by zinc and cadmium in solutions of their sulphates, whilst silver, lead and mercury, in solutions of the corresponding nitrates, give perfect straight lines with no visible irregularity, S.P. 1 and S.P. 2 being identical, and exactly in line with the main curve. A table of results with these electrodes would therefore consist of a series of zeros and need not be given.

The obvious conclusion from this work is that if a metal only is being deposited, there is no overvoltage and no transfer resistance—in other terms, metal overvoltage does not exist.

When similar experiments are tried with iron, nickel or cobalt in solutions of the corresponding sulphates, chlorides or nitrates, the curves obtained are

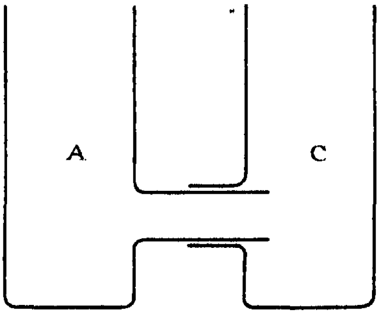


FIG. 1.

FIG. 2.

FIG. 3.

(*Facing p.* 682.)

entirely different in character—S.P. 1 and S.P. 2 no longer coincide, and it is impossible to deposit metal without liberating hydrogen at the same time.

Fig. 3 illustrates the type of curve given by these metals, and was obtained with a nickel cathode in a normal solution of nickel sulphate, the current density being 10 milliamperes per square centimetre. The close resemblance of these curves to those obtained in pure acid electrolyte, together with the fact that hydrogen is being liberated, at once suggests that these are really hydrogen overvoltage curves, and this is confirmed by further study of the data obtained. In the following table, therefore, overvoltages are calculated from the hydrogen electrode as standard. The single potentials given are based on the assumption that the single potential of the $Hg_2SO_4 . N\ H_2SO_4$ is + 0·97 volt.

Cathode.	Electrolyte.	S.P. I.	S.P. 2.	C.D.	O.V.	T.R.
Iron	$FeSO_4$	0·0	− 0·13	1	0·65	25
„	„	0·0	− 0·13	10	0·65	8
„	$Fe_2(SO_4)_3$............	− 0·16	− 0·16	100	0·5	0·5
Nickel	$NiSO_4$	+ 0·55	+ 0·02	1	0·7	50
„	„	+ 0·55	− 0·04	10	0·7	18
Cobalt	$CoSO_4$	+ 0·24	+ 0·07	1	0·3	30
„	„	+ 0·24	+ 0·07	10	0·3	8

S.P. 1.—Single potential before use in volts. S.P. 2.—Single potential after use in volts.
C.D.—Current density in milliamperes per square centimetre.
O.V.—Hydrogen overvoltage in volts. T.R.—Transfer resistance in ohms per square centimetre.

Discussion of Results.

The difference between the behaviour of the metals of the iron group from that of all the other metals is very striking, and the reason for this difference is obscure. The author has suggested* that it may be due to the ions of these metals being hydrated and the ions of other metals non-hydrated, but the evidence in favour of this theory is not very convincing. Glasstone† has suggested that these three metals are deposited with an unstable electronic arrangement which slowly reverts to the stable form. This theory is very attractive, and would assist in the explanation of many anomalies, but a vital objection arises from the fact that the same values for the overvoltage may be obtained, whether the metal in question is being deposited or not. Thus a nickel cathode shows the same overvoltage when the nickel sulphate solution is replaced by sodium sulphate solution; a cobalt cathode shows the same overvoltage in dilute sulphuric acid, sodium hydroxide, cobalt sulphate, cobalt chloride, etc. The one condition necessary to obtain these values is not the deposition of metal but the liberation of hydrogen, and one is therefore compelled, somewhat reluctantly, to reject Glasstone's very ingenious theory, at least in its present form. A further point which is difficult to reconcile with Glasstone's theory lies in the sudden large changes of overvoltage which occur when these metals are being deposited. Such changes are common when measuring hydrogen, oxygen or chlorine overvoltages, and are readily explainable on the basis of the overvoltage-compound or valency theory. It is remarkable that iron in ferrous sulphate and nickel in nickel sulphate solutions both show the high overvoltage characteristic of a divalent metal, whereas iron in ferric sulphate gives a value 0·5 volt characteristic of a trivalent metal. This value is only obtained at comparatively high current densities, since at low current densities all the hydrogen is used in reducing the solution and there is no overvoltage and no transfer resistance. A cobalt cathode in cobalt sulphate solution generally shows an overvoltage of about 0·3 volt, but if a fairly high current density is applied for some time this value rises rapidly to over 0·6 volt. In a cobalt nitrate solution this sudden rise occurs more readily, and the overvoltage (0·7 volt) is then characteristic of a divalent metal.

---

* 'J. Chem. Soc.,' *loc. cit.*

† 'J. Chem. Soc.,' vol. 129, p. 2896 (1926).

Many further similar features could be quoted, but sufficient has been said to show that the overvoltages shown by the three metals, iron, nickel and cobalt, during metal deposition are hydrogen overvoltages, and are therefore due to the high single potentials of metallic hydrides formed on or in the electrode surface. These single potentials are not governed by the hydrogen-ion concentration of the electrolyte any more, for example, than that of a copper electrode in a solution of a copper salt is governed by the pH value of the solution. The hydrides do not give rise to hydrogen ions, but ionise themselves *in toto* by simple loss of an electron, exactly as a metal behaves. Glasstone's objection to the hydride theory (*loc. cit.*) therefore appears to be founded on a misapprehension.

We are still, however, without an explanation of the reason why these three metals behave in so different a manner from all the other common metals, and here it may be possible to retain a portion of Glasstone's theory.

The chief points requiring explanation are (*a*) that there appears to be some peculiar force acting in these three cases tending to strongly retard the separation of these metals from solution, and (*b*) there is an extraordinary difficulty in determining the true normal single potentials of these metals, values differing by as much as 0·87 volt being obtained by different workers in the case of nickel alone.

As a consequence of (*a*) it is impossible to deposit these metals electrolytically in the pure state, hydrogen being the usual impurity present. If a strongly electro-negative metal such as zinc is present, at high current densities very little iron group metal or hydrogen is deposited even in an acid solution until most of the zinc is exhausted, and this appears to indicate that hydrogen also is subject to a similar retarding force. If, therefore, Glasstone's theory is applied to iron, nickel and cobalt, it must also be applied to hydrogen, and this presents very grave difficulties. The energy changes associated with the change of the one electron of the hydrogen atom from one orbit to another are far greater than those postulated by Glasstone for the corresponding case of nickel, and the problem is still further complicated by the fact that the hydrogen ion has no electrons. It is perhaps possible to devise a theory whereby the electron received by the hydrogen ion from the electrode enters a different orbit from that which it occupies in the normal state, and to base a new theory of over-voltage upon this; but such a theory appears to the author to be too speculative and improbable.

The alternative theory suggested by the author, that the ions of these three metals and of hydrogen are hydrated, is open to less objection than that of Glasstone. Unfortunately the whole question of hydration of ions is still very controversial, but the evidence in favour of the hydration of the hydrogen ion appears to be very strong.* It is therefore probable, though not definitely proved, that the ions of iron, nickel and cobalt are also hydrated, and that the retarding force in question is due to the accumulation of hydrated discharged ions on the electrode surface, time being required by these atoms to get rid of the water of hydration. This theory has already been discussed by the author,† and it has been shown that the behaviour giving rise to (*b*) above may also be accounted for on the same theory; but in this connection the action of the electrolyte on the electrode surface, with formation of a film of insoluble or colloidal hydroxide, introduces further complications. On the whole, therefore, we may conclude that whereas Glasstone's theory appears very improbable, the hydrated ion theory reasonably accounts for the observed facts, although definite proof of its truth is not yet forthcoming.

---

* Lapworth, 'J. Chem. Soc.,' vol. 107, p. 857 (1915).

† *Loc. cit.*

Summary.

From experiments on metal deposition with the aid of the cathode ray oscillograph, it is shown that overvoltage and transfer resistance do not occur unless a gas is being liberated at the electrode.

Hydrogen is always deposited along with metal when the metals of the iron group are separated electrolytically from pure solutions of their respective salts, and this gives rise to hydrogen overvoltage at the cathode.

The reason why the metals of the iron group show this exceptional behaviour is discussed, and it is shown that it is probably due to hydration of the ions of these metals.

A defect has been discovered in the design of the cell previously used for these experiments. This defect is responsible for erroneous values for transfer resistance, but has no effect upon the values for overvoltage.

A Revision of the Theory of Transfer Resistance.

By Edgar Newbery, D.Sc., F.I.C., University of Cape Town.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 20, 1928.)

In the preceding paper on " Metal Overvoltage " it has been shown that a defect in the design of the electrolytic cell used in previous work gave rise to erroneous values for transfer resistance. It is necessary, therefore, to examine to what extent, if any, the theory of transfer resistance must be modified in order to adapt itself to the new values obtained when the above defect is remedied.

A number of experiments have been repeated with the various electrodes, using the new type of cell, but since the conclusions arrived at were the same in all cases, it will be sufficient to give results for a single electrode. The following values were obtained with a silver cathode in an electrolyte of N sulphuric acid at room temperature.

Current density.	Transfer resistance. Old cell.	Transfer resistance. New cell.
Milliamperes.	Ohms.	Ohms.
1	35	40
10	16	8·5
100	5·7	1·5
1000	0·8	0·3

At the lower current densities where the transfer resistance is high, the change of cell makes little difference. With silver the difference appears marked at a C.D. of 10 milliamperes per square centimetre, but with other metals very little difference is observed until higher current densities are applied.

At higher current densities there is no doubt that the old type of cell gave values for transfer resistance which were much too high, and this effect is still more marked if electrolytes of lower conductivity are used. Thus copper in copper sulphate solution gave an apparent value of about 15 ohms at all current densities, whereas with the new cell the value was nearly zero.

Perhaps the most important effect of using the new cell for these measurements appears when the concentration of the electrolyte is changed. In a previous communication* a table showing this effect was given. This table is undoubtedly incorrect. A series of experiments was tried with the new cell, using acid varying in concentration between 0·1 normal and 10·0 normal, and current densities from 1 to 1000 milliamperes per square centimetre, and identical results at all concentrations were obtained. This result was further confirmed using a copper cathode in place of a silver one. It is evident, therefore, that transfer resistance is quite independent of the concentration of the electrolyte.

The effect of temperature on transfer resistance is also shown by the new cell more correctly than by the old one. Previous experiments gave discordant results, in some cases positive and in others negative, but always small. With the new cell, however, quite definite increases of transfer resistance are observed, with increasing temperature. The following table illustrates this :—

Electrode.	Current density.	T.R. 15° C.	T.R. 50° C.
Copper	10	6·6	10
,, ...	100	1·4	2·0
Silver	100	1·5	1·7

---

* 'Roy. Soc. Proc.,' A, vol. 111, p. 188 (1926).

It is evident that the discordant results obtained with the older cell were due to the negative temperature coefficient of the resistance of the electrolyte interfering with the positive temperature coefficient of the transfer resistance.*

All the above data point to the conclusion that in the new type of cell the disturbing influence of the fall of potential through the electrolyte has been effectively eliminated, and we may therefore summarise the known facts about transfer resistance as follows :—

(1) Transfer resistance occurs only when a gas is being liberated at an electrode.

(2) The highest values are obtained with very low current densities. At high current densities transfer resistances become very small.

(3) Smooth and polished surfaces give higher values than rough or spongy surfaces.

(4) The conductivity of the electrolyte has no influence on transfer resistance.

(5) Transfer resistance increases with rise of temperature.

Theory of Transfer Resistance.

In the previous discussion on transfer resistance it was suggested that it was due to two factors : (*a*) the resistance of a film of gas over the electrode surface, and (*b*) the resistance of a layer of partially exhausted electrolyte surrounding the electrode.

In view of (1) and (4) above, the second part of this suggestion must now be withdrawn, and the first factor alone will then be responsible for the whole.

The diminution at high current densities is due partially to the compression of this film of gas by the extremely high mechanical pressure between the electrically charged electrode and the oppositely charged ions on the other side of the very thin layer of gas, and partly to the very high potential gradient enabling electrons to shoot right through the film. The lower values at rough surfaces are due to the greater facilities for escape of gas, and therefore greater difficulty in forming a continuous gas film.

The positive temperature coefficient is probably due to the increase in thickness of the gas film by heat and consequent increase of resistance.

This increase of thickness is not necessarily due to simple thermal expansion of the gas. The lowering of the surface tension of the electrolyte by rise of temperature will allow a thicker film of gas to form before it can escape in the form of bubbles. This variation of transfer resistance with surface tension has led to attempts by certain workers to explain overvoltage phenomena in terms of surface tension, bubble diameter, etc. The small increase of transfer resistance produced by the addition of certain colloids is probably due to the same cause.

---

* It may be noted here that the values obtained for the overvoltages of all the metals are quite unaffected by the type of the cell used, and hence the theory of overvoltage already put forward stands unaltered. It may also be mentioned that in the course of this work it was found that silver shows two definite hydrogen overvoltages, 0·0 volt and 0·35 volt, but the exact conditions under which either extreme is shown were not fully discovered. The usual values obtained lie between these extremes.

We may conclude, therefore, that transfer resistance is the electrical resistance of a film of gas covering the surface of an electrode.

On the Kinetic Method in the New Statistics and its Application in the Electron Theory of Conductivity.

By Dr. L. W. Nordheim.

(Communicated by R. H. Fowler, F.R.S.—Received May 30, 1928.)

Introduction and Summary.

Hitherto in work on the new statistics the method of probabilities has always been used, a method which consists in the measuring of the phase space by counting complexions. Now this method is certainly very useful and valuable on account of its generality and convenience, but it has the drawback that only equilibrium states and fluctuations can be treated, not the more general streaming and conductivity phenomena. In the classical physics a method for such problems has been created in the dynamical theory of gases by Maxwell and Boltzmann, and it seems therefore useful to try to adapt it to the new statistics.

That task has become particularly necessary, because a corresponding method, due to H. A. Lorentz, has been applied by Sommerfeld* in his theory of metallic conductivity. As it will be shown, the dynamical theory can be worked out in the new statistics, both for the Einstein-Bose and the Fermi-Dirac statistics, in a way completely analogous to the corresponding one of Boltzmann, only some characteristic additional terms occurring. The latter, however, just cancel out in the special case of the electronic conductivity under certain assumptions, so that Sommerfeld's procedure is justified.

Therefore in §§ 1 and 2 the general theory of gases will be developed, and in § 3 it will be applied to the theory of conductivity.

---

* Sommerfeld, 'Z. f. Physik,' vol. 47, p. 1 (1928).

§ 1. *The Fundamental Equation of the Theory of Gases.*

We consider, as usual, an assembly of N mass points of mass m, which may be subjected to any exterior fields of force but interact with each other only at very small distances. The macroscopic state of such an assembly is then given in the well-known manner by a distribution function

$$f(x, y, z, \xi, \eta, \zeta),$$

which gives us the number of particles in every infinitesimal range $\Delta x\, \Delta y\, \Delta z\, \Delta\xi\, \Delta\eta\, \Delta\zeta$ (abbreviated to $\Delta V\, \Delta v = \Delta\Omega$) of the phase-space and which is so normalised that

$$\int f\, \Delta\Omega = N, \tag{1}$$

so that the integration over any part of the phase space gives directly the number of particles in that part. Furthermore, we shall use throughout the convention, that in other than Cartesian co-ordinates and momenta $\Delta\Omega$, or ΔV and Δv, include the functional determinant for the co-ordinates used, so that f is an invariant against all transformations of that kind. It gives simply for each point of the phase space the number of particles in the unit element $\Delta\Omega$ constructed around that point.

According to the fundamental principles of the quantum theory such a description is permissible only if the elements $\Delta\Omega$ are supposed to be large compared with h^3/m^3. As formerly a similar convention was necessary, which naturally still has to be made, *i.e.*, that any such element still contains many particles and that these can be regarded as evenly distributed (hypothesis of molecular chaos), the above restriction does not need to disturb us.

The classical procedure consisted, as is well known, in the computation of the total time rate of alteration of the distribution and equating it to zero as the condition of stationary flow, and the same treatment will be followed here. The variation in time consists now of two parts, *i.e.*, of the variation by the streaming of the points representing the particles in the phase space according to their own motion and the accelerations by external forces, and, secondly, by the collision-like interactions between the particles themselves. Thus, taking a phase element $\Delta\Omega$, the Boltzmann fundamental equation becomes

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\xi + \frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial z}\zeta + \frac{\partial f}{\partial \xi}\frac{X}{m} + \frac{\partial f}{\partial \eta}\frac{Y}{m} + \frac{\partial f}{\partial \zeta}\frac{Z}{m} + a - b = 0. \tag{2}$$

Here X, Y, Z are the components of the external force and a, b respectively the number of particles thrown out of or into the element $\Delta\Omega$ by impacts in unit time.

The flux terms apparently remain unaltered by the quantum theory. This follows from the considerations of Darwin* or Kennard,† who have shown that a single undisturbed particle, which can be represented in the quantum mechanics by a wave packet, behaves itself under the influence of external electric or magnetic forces very approximately as in the classical mechanics. If there are several particles, the interaction of which can be neglected, then each of them is to be represented by a wave packet, and such packets can move straight through each other, *i.e.*, during a " collision," when the centres of the packets are not farther away from each other than their extension, a complicated interference phenomenon occurs, but afterwards the particles separate according to ordinary mechanics, or in the quantum theory the wave groups clear themselves from each other quite distinctly without further interference. The prohibition of certain states of motion, characteristic for quantum statistics, emerges here not out of the law of motion but out of the choice of the initial state, *i.e.*, of a proper wave group. Thus, for instance, it is prohibited in the Fermi-Dirac statistics to choose an initial distribution, the density in the phase space of which is greater than one particle per h^3/m^3.

With the collision term, however, it is another matter. In classical physics the probability of a collision-interaction of two particles depends only on the number of particles in the initial states, but in quantum theory it depends on the number in the final states too. The corresponding laws have already been developed by Jordan,‡ Ornstein and Kramers,§ and most generally by Bothe,|| and we have only to formulate them appropriately.

Firstly, we have to define the " collision processes " in a suitable way for the quantum mechanics. In the classical theory a collision process is entirely determined by the position and velocity of the particles, and accordingly exact variables for the impacts are introduced (for instance, the parameter, *i.e.*, the smallest distance to which the particles would approach in rectilinear motion). In quantum mechanics that method is inadmissible on account of the uncertainty relation, and one has to proceed instead in the following way: We take a volume element ΔV, which shall be large compared with the " atomic dimensions." Through that element two different proper normalised wave groups, representing particles with velocity vectors between $\mathbf{v}$ and $\mathbf{v} + \Delta\mathbf{v}$

---

* Darwin, 'Roy. Soc. Proc.,' A, vol. 117, p. 258 (1927).

† Kennard, 'Z. f. Physik,' vol. 44, p. 326 (1927).

‡ Jordan, 'Z. f. Physik,' vol. 41, p. 711 (1927).

§ Ornstein and Kramers, 'Z. f. Physik,' vol. 42, p. 481 (1927), who suggest already an analogous procedure to that of this paper without carrying it through.

|| Bothe, 'Z. f. Physik,' vol. 46, p. 327 (1928).

and $\mathbf{v}_1$ and $\mathbf{v}_1 + \Delta\mathbf{v}$ (of course $\Delta V\,\Delta\mathbf{v}$ has to be $\gg h^3/m^3$), may run. Then a probability $\phi^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*}$ exists that a scattering into the velocity vector ranges $\mathbf{v}^*$ to $\mathbf{v}^* + \Delta\mathbf{v}$, and $\mathbf{v}_1^*$ to $\mathbf{v}_1^* + \Delta\mathbf{v}$ may occur. Of course $\mathbf{v}^*$ and $\mathbf{v}_1^*$ are not entirely free but have to obey the energy and momentum theorems, *i.e.*, for given initial values $\mathbf{v}$ and $\mathbf{v}_1$ they have to lie on an element $d\sigma$ of the corresponding energy-momentum surface in the phase space. We shall use therefore also the notation $\phi^{\mathbf{v}\mathbf{v}_1}_{(\sigma)}$ to describe the probability. $\phi_{(\sigma)}$ could be calculated in the classical mechanics from the assumed law of force and the chaos hypothesis, and similarly in the quantum mechanics from the assumed particle model.

In the old statistics, then, the probability of such a process in an assembly of particles depends only on the number of particles in the initial states before the process, that is, in an obvious notation

$$P^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*} = f f_1 \phi^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*}\,\Delta V\,\Delta\mathbf{v}\,\Delta\mathbf{v}_1. \tag{3}$$

In this formula f and f_1 denote the value of the distribution function for the same space point but the different velocities $\mathbf{v}$ or $\mathbf{v}_1$.

In the quantum theory, however, the probability depends also on the number of atoms already in the final states. That is the fundamental discriminating point, and the two quantum statistics are also here differentiated.

For the Fermi-Dirac statistics the essential feature is the impossibility of more than one particle having room in a phase element of the magnitude $h^3/(m^3 g)$ (g being the statistical weight of one quantum state). If, therefore, such a cell is already occupied, then a process in which another particle would be thrown into it cannot occur. If the phase space is not completely dense packed, correspondingly only a reduction of the probability takes place in the ratio of the still available to the total room. Thus, if in the unit volume of the phase space A quantum cells exist (for the ordinary translation motion A is, of course, equal $m^3 g/h^3$, and according to our definition of $\Delta\Omega$ it is invariant against transformations of the co-ordinates or momenta), the probability of a process, as explained above, becomes in the Fermi-Dirac statistics

$$P^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*} = f\left(1 - \frac{f^*}{A}\right) f_1 \left(1 - \frac{f_1^*}{A_1}\right) \phi^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*}\,\Delta V\,\Delta\mathbf{v}\,\Delta\mathbf{v}_1. \tag{4}$$

In the Einstein-Bose statistics, on the other side, the probability is *increased* by the existence of other particles (that is, for instance, the meaning of the Einstein coefficient of induced radiation), and one has therefore in that case

$$P^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*} = f\left(1 + \frac{f^*}{A}\right) f_1 \left(1 + \frac{f_1^*}{A_1}\right) \phi^{\mathbf{v}\mathbf{v}_1}_{\mathbf{v}^*\mathbf{v}_1^*}\,\Delta V\,\Delta\mathbf{v}\,\Delta\mathbf{v}_1. \tag{5}$$

With these expressions the difference $a - b$ of the number of particles thrown out of or into the phase element $\Delta\Omega$ can now be formed in the well-known way, and one gets from (2) and (4) or (5) the fundamental equation

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\xi + \frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial z}\zeta + \frac{\partial f}{\partial \xi}\frac{X}{m} + \frac{\partial f}{\partial \eta}\frac{Y}{m} + \frac{\partial f}{\partial \zeta}\frac{Z}{m}$$

$$= -\iint\left[f\left(1 \pm \frac{f^*}{A}\right)f_1\left(1 \pm \frac{f_1^*}{A_1}\right) - f^*\left(1 \pm \frac{f}{A}\right)f_1^*\left(1 \pm \frac{f_1}{A_1}\right)\right]\Delta v_1 \int \phi_{(\sigma)}^{v\, v_1}\, d\sigma. \tag{7}$$

It is here assumed, as usual, that inverse processes have to be equally probable, *i.e.*, that

$$\phi_{v^*\, v_1^*}^{v\, v_1} = \phi_{v\, v_1}^{v^*\, v^*_1}. \tag{8}$$

That relation in the classical mechanics asserts the reversibility of all mechanical processes. In the quantum mechanics it is guaranteed by the Hermitian condition

$$F_{v^*\, v_1^*}^{v\, v_1} = \bar{F}_{v\, v_1}^{v^*\, v^*_1}$$

for the transition matrix

$$F_{v^*\, v_1^*}^{v\, v_1} = \int \psi\,(v\ v_1)\; F\bar{\psi}(v^*\ v_1^*)\; dV\, dV_1,$$

where $\psi\,(v\ v_1)$ is the proper normalised eigenfunction for the state of movement v, v_1, and F the operator corresponding to the interaction energy between the particles, and the transition probabilities are given by

$$\Phi_{v^*\, v_1^*}^{v\, v_1} = |\, F_{v^*\, v_1^*}^{v\, v_1}\,|^2. \tag{9}$$

In equation (7) the upper sign (+) is valid for the Einstein-Bose, the lower (−) for the Fermi-Dirac statistics, whilst for the classical theory these additional terms are to be omitted altogether. We return to the classical theory in the limiting case of a density so small everywhere that throughout the phase space $f/A \ll 1$.

Instead of the Cartesian co-ordinates and momenta, of course also other suitable co-ordinates can be introduced, for instance, energy and direction of motion.

The treatment of mixtures of two or more kinds of particles offers no difficulties, even if they obey different statistics, and it seems hardly necessary to go here farther into the details. Then collision terms for the new interaction possibilities are to be added, which have accurately the same structure as the right side of (7), f_1, f_1^* and A_1 only referring to another kind of particle.

§ 2. *Equilibrium States and the H-Theorem.*

We have now to prove that the fundamental equation (7) allows us actually to build up a complete gas theory, and leads especially to the known distribution functions of the new statistics. Again, the deduction is exactly analogous to the old one of Boltzmann.

In the equilibrium case with no external forces the left side of (7) vanishes, and therefore the integral on the right side has to be zero. As it will be proved later, the necessary and sufficient condition for this is the vanishing of the square bracket, *i.e.*, the relation

$$f\left(1 \pm \frac{f^*}{A}\right) f_1\left(1 \pm \frac{f_1^*}{A_1}\right) = f^*\left(1 \pm \frac{f}{A}\right) f_1^*\left(1 \pm \frac{f_1}{A_1}\right). \tag{10}$$

If we introduce the new function g instead of f by the definition

$$f = \frac{A}{g \mp 1},$$

which implies no restriction of generality, as f is a one-valued function of g, we obtain from (10)

$$g^* g_1^* = g\, g_1. \tag{11}$$

Since f and therefore also g can only depend on the modulus but not the direction of the vector of velocity, and since, further, the energy relation

$$|\mathrm{v}_1^*|^2 = |\mathrm{v}|^2 + |\mathrm{v}_1|^2 - |\mathrm{v}^*|^2$$

holds between the four velocity vectors, it follows as in Boltzmann's argument that

$$g = Ce^{c|\mathrm{v}|^2},$$

where C and c are constants. Therefore

$$f = \frac{A}{Ce^{c|\mathrm{v}|^2} \mp 1}, \tag{12}$$

i.e., the well-known Einstein or Dirac formula.

It remains to prove the necessity of (10). For that purpose we establish the H-theorem. Multiplying the fundamental equation (7) by $\log \frac{f}{A \pm f} \Delta\Omega$ and integrating over the whole phase space, the left side becomes the total differential quotient with respect to the time of the expression

$$H = \int [f \log f \mp (A \pm f) \log (A \pm f)]\, \Delta\Omega, \tag{13}$$

since

$$\frac{dH}{dt} = \int \frac{d}{dt}[f \log f \mp (A \pm f) \log (A \pm f)]\, \Delta\Omega$$
$$= \int [\log f - \log (A \pm f)] \frac{df}{dt}\, \Delta\Omega.$$

On the right side one gets firstly

$$\frac{d\mathrm{H}}{dt} = -\iint \log\frac{f}{\mathrm{A} \pm f}[f(\mathrm{A} \pm f^*) f_1 (\mathrm{A}_1 \pm f_1^*) - f^* (\mathrm{A} \pm f) f_1^* (\mathrm{A}_1 \pm f_1)] \\ \times \frac{1}{\mathrm{A}\,\mathrm{A}_1} \Delta\mathrm{V}\, \Delta\mathrm{v}\, \Delta\mathrm{v}_1 \int \phi(\sigma)\, d\sigma.$$

This relation must be unaltered by an interchange of f and f_1, which means only an alteration in the notation, so that one can write

$$\tfrac{1}{2}\left(\log\frac{f}{\mathrm{A} \pm f} + \log\frac{f_1}{\mathrm{A}_1 \pm f_1}\right) \quad \text{instead of} \quad \log\frac{f}{\mathrm{A} \pm f}$$

under the integral. It must also be unaltered by an interchange of the functions with stars against those without stars and vice versa (which corresponds only to an inverting of all collision processes), so that this may be replaced (the square bracket changing its sign by the said procedure) by

$$\tfrac{1}{4}\left(\log\frac{f}{\mathrm{A} \pm f} + \log\frac{f_1}{\mathrm{A}_1 \pm f_1} - \log\frac{f^*}{\mathrm{A} \pm f^*} - \log\frac{f_1^*}{\mathrm{A}_1 \pm f_1^*}\right) \\ = \tfrac{1}{4}\{\log[f(\mathrm{A} \pm f^*) f_1 (\mathrm{A}_1 \pm f_1^*)] - \log[f^* (\mathrm{A} \pm f) f_1^* (\mathrm{A}_1 \pm f_1)]\}.$$

One has therefore the relation

$$\frac{d\mathrm{H}}{dt} = -\tfrac{1}{4}\iint \{\log[f(\mathrm{A} \pm f^*) f_1 (\mathrm{A}_1 \pm f_1^*)] - \log f^* (\mathrm{A} \pm f) f_1^* (\mathrm{A}_1 \pm f_1)]\} \\ \times \{f(\mathrm{A} \pm f^*) f_1 (\mathrm{A}_1 \pm f_1^*) \\ - f^* (\mathrm{A} \pm f) f_1^* (\mathrm{A}_1 \pm f)\} \frac{\Delta\mathrm{V}\, \Delta\mathrm{v}\, \Delta\mathrm{v}_1}{\mathrm{A}\,\mathrm{A}_1} \int \phi(\sigma)\, d\sigma. \quad (14)$$

Now the integral is always positive, since both brackets have everywhere the same sign. Therefore H can only decrease, and the expression (13) appears thus as the complete analogon of the H-function of Boltzmann. In the equilibrium state especially $d\mathrm{H}/dt$ has to be zero, which is only possible if the relation (10) holds.

In comparison with the well-known results of the Fermi-Dirac and Einstein-Bose statistics one sees that indeed according to (13)

$$\mathrm{H} = -\log \mathrm{P}, \qquad (15)$$

where P is the " complexion number " calculated in the usual way, the integral in (13) instead of the common summation being only a formal difference. $-k\mathrm{H}$ gives, therefore, the correct value for the entropy, and thus the same thermodynamic results as the statistical method.

From the preceding considerations it seems quite clear that the entire theory of gases can be developed in a completely analogous way as in the classical physics, and that especially equation (7) gives the right generalisation of the fundamental equation. It would perhaps not be too difficult to carry those calculations farther on, and, for instance, to compute the viscosity and conduction of heat at least in the case of nearly complete degeneration. But as at the present time a possibility of observation of such effects hardly seems to exist, we shall not go into a further discussion here.

§ 3. *On the Electron Theory of Metals.*

The deviation from the classical laws, being of importance for ordinary gases only under very extreme conditions, plays, however, the fundamental rôle in the electron theory of metals according to Sommerfeld, and therefore his procedure seems at first somewhat inconsequent, as he uses the new distribution law, but also the old fundamental equation (in the special form of H. A. Lorentz), which would of itself lead only to the old formula, and the present discussion has been undertaken particularly to discuss that question.

In the electron theory of conductivity, which consists in assuming a great many dissociated electrons forming something like an electron gas inside the metals, the collisions of the electrons with the metal ions are the more important, since by them the free paths, during which the external electric forces are able to accelerate the electrons undisturbed, are interrupted, whilst at an impact of electrons with each other the momentum gained by one electron is transferred to the next one.

Now the best way of describing the interactions of the electrons with the ion lattice is apparently again the introduction of certain scattering coefficients $\psi_{\mathbf{v}^*}^{\mathbf{v}}$, which ascribe to every electron with the vector velocity $\mathbf{v}$ a probability of being scattered into the new state $\mathbf{v}^*$.

These $\psi_{\mathbf{v}^*}^{\mathbf{v}}$ depend, of course, on the state of the metal, *i.e.*, on its temperature and so on. Without going into further details of investigation of those expressions, we can establish the collision term as follows, the electrons obeying the Fermi-Dirac statistics,

$$a - b = \int\left[f\left(1-\frac{f^*}{A}\right)\psi_{\mathbf{v}^*}^{\mathbf{v}} - f^*\left(1-\frac{f}{A}\right)\psi_{\mathbf{v}}^{\mathbf{v}^*}\right]\Delta\mathbf{v}^*. \tag{16}$$

From this we see now that the additional terms due to the new statistics just cancel out if only

$$\psi_{\mathbf{v}^*}^{\mathbf{v}} = \psi_{\mathbf{v}}^{\mathbf{v}^*}. \tag{17}$$

But that relation is now, in contrast to (8), by no means necessary. The scattering medium itself has to be conceived as a discontinuum in some statistical equilibrium state. Equation (16) has therefore to be understood as an abbreviated writing of (7) (whereby f_1 and f_1^* and A_1 and A_1^* refer, of course, to the state of the scattering medium), and (17) is therefore valid only if the relation

$$f_1\left(1 \pm \frac{f_1^*}{A_1}\right) = f_1^*\left(1 \pm \frac{f_1}{A_1}\right),$$

i.e., $f_1 = f_1^*$ holds. This condition means now apparently that the scattering system is not influenced noticeably by the interaction with the electrons, and that, in fact, is the assumption usually made in the electron theory of conductivity. It is justified by the large ratio of the masses of the atoms to those of the electrons, which has the result that no energy is transferred between them. This being so, the collisions between electrons and metal have no influence on the equilibrium distribution, which is effected only by the interaction between the electrons themselves, while, on the other side, these collisions are of no importance for the computation of the conductivity, at least to a first approximation.

Thus, although by the dense packing of the electrons the frequency of scattering processes is reduced very much, the difference in the numbers of electrons thrown out of and into a certain bundle, which alone is responsible for the resistance, remains unaltered. We are therefore justified in this special case in proceeding as if there were no degeneration at all. Sommerfeld has already completely worked out the theory from that point of view, so that there remains only the determination of the scattering probability, *i.e.*, the "free path" of the electrons, as the chief problem. That problem has been attacked in a most ingenious way by Houston,* who compares the scattering of electrons in a metal lattice with the scattering of X-rays. This method seems to be quite suitable, since both the Schrödinger waves of the electrons and the electromagnetic waves obey the same partial differential equation $\nabla^2\psi - \kappa\psi = 0$ (of course after removing the time factor). That method itself gives naturally only a correct solution for a wave representing a single electron, as in the collision theory of Born, whilst really the wave equation should have three degrees of freedom for each electron. But by the analysis of this paper it is generally established that in the case of electric conductivity the general solution for many electrons may be replaced by the special solution for a single one only.

---

* Houston, 'Z. f. Physik,' vol. 48, p. 449 (1928). I am very much indebted to Mr. Houston for allowing me to see his work before publication.

Finally, I wish to express my best thanks to Dr. R. H. Fowler, F.R.S., for his interest and assistance in the writing of this paper and for many valuable discussions about the electron theory of metals, and to the International Education Board for enabling me to work at Cambridge.

The Fluorescence of Mercury Vapour.

By R. W. Wood, For. Mem. R.S., and V. Voss.

(Received June 5, 1928.)

[Plate 14.]

The recent important and very interesting investigations by Lord Rayleigh of the varied phenomena accompanying the production of the band spectra of mercury by the electrical discharge and optical excitation have added much to our knowledge regarding the emission spectra of this element in the molecular condition, and have shown that the mechanism of emission is even more complicated than previously supposed. Lord Rayleigh's investigations have concerned the after-glow of the electrically excited vapour for the most part, and it appeared to be worth while to renew the attack upon the fluorescence excited by the light of the aluminium spark, having in mind his recent discoveries.

In the present paper we shall show that practically all of the stronger arc lines of the mercury atom are emitted by mercury vapour when excited by the light of the aluminium spark, which can be absorbed only by the mercury molecule (band absorption). The intensity of these lines varies as the *square* of the intensity of the exciting light, as was shown recently by one of us in the case of *atomic* mercury excited by a two-stage absorption process. The effects of superheating the vapour on the intensities of the various lines and bands has been investigated, as well as the development of other bands and lines due to the presence of impurities, and caused probably by collisions with excited molecules.

A recent paper by Niewodniczanski has also reopened the question as to why, *under certain conditions*, the vapour exhibits fluorescence only when undergoing distillation, as was proved by an experiment devised by one of the present writers, and described in papers published in 1921.* At that time experimental evidence was conflicting, as mercury appeared at times to fluoresce when stagnant, and at other times only when in the process of distillation. The crucial experiment above referred to, which at the time appeared to have settled the matter, was made with two small exhausted quartz bulbs joined by a short narrow tube, and containing a drop of mercury. These were heated in a current of hot air rising from an asbestos chimney surrounded by a nest of bunsen burners, and illuminated by the light of an aluminium spark. With both bulbs at the same temperature (210°) no fluorescence was seen, but if a blast of cold air was directed against one of the bulbs, the other immediately burst into brilliant green fluorescence, which continued until all of the mercury had condensed in the cooled bulb, when it immediately faded away.

Niewodniczanski challenged this experiment without, however, trying it apparently, claiming that he always obtained fluorescence with uniformly heated tubes, and that cooling a part of the system invariably decreased the fluorescence.

It was with a view of finding out the cause of this conflicting evidence that the present investigation was commenced. It became apparent almost immediately that the behaviour of bulbs depended very much upon how carefully they had been freed from occluded water vapour before sealing from the pump, and we presently were able to seal off a bulb from the pump and bring it into a condition in which it would show fluorescence only when the mercury was distilling, or into a condition in which it fluoresced at constant temperature with no distillation going on. The bulb contained a trace of water vapour, which could be driven back into the quartz by a high-frequency electrodeless discharge, the bulb becoming almost non-conducting. After this treatment it showed brilliant fluorescence at constant temperature. A small flame was now played over the surface, liberating the occluded water vapour, and fluorescence now occurred only when distillation took place, *i.e.*, with unequal heating or local cooling. The bulb was carried back and forth from one condition to the other for a number of times. The mystery now disappeared entirely. Distilling mercury vapour drives the water vapour ahead of it, the pure mercury vapour freshly formed and unmixed with the impurity showing fluorescence. We prepared one of the double bulbs with connecting tube, and heated it for 2 hours to a high temperature, the pump operating on it all the time; mercury was then distilled into them and they were sealed off. No distillation was necessary in this case, both bulbs showing brilliant fluorescence in the hot-air column. After prolonged operation, however, a trace of water vapour developed and the fluorescence disappeared, and could be developed only by cooling one bulb; mercury vapour then passed over from the hot to the cooled bulb, sweeping the water vapour along with it, and the fluorescence developed in the freshly formed pure vapour. This result was reported at the conference in Physics held in the Loomis Laboratory at Tuxedo, N.Y., on January 6, and in 'Nature,' February, 1928.* It has been confirmed by work carried out independently by Pringsheim and Terenin, and published at about the same time in the 'Zeitschrift fur Physik.'†

---

* R. W. Wood, 'Roy. Soc. Proc.,' A, vol. 92, p. 362 (1921); van der Lingen and Wood, 'Astro-Phys. J.,' vol. 54, p. 149 (1921).

As illustrating the baffling nature of the results which we obtained at first, the following experiments may be cited: We prepared bulbs both single and double which we baked out on the pump and, after admitting a little air, exhausted to pressures varying from 0·001 mm. to 0·4 mm. They all showed bright fluorescence at constant temperature of 250° except the one at the lowest pressure (10^{-4} mm.), which exhibited only a very feeble fluorescence, which became bright, however, as soon as local cooling was applied. The inference from this experiment, taken by itself, might be that the presence of a small amount of air was necessary to cause fluorescence with stagnant vapour, and that distillation was necessary in a more perfect vacuum, a most improbable circumstance. We next found that if the bulb was more thoroughly baked out and the air admitted through a long column of P_2O_5, that fluorescence was obtained with stagnant vapour at all air pressures between zero and 0·4 mm. After numerous control experiments under modified conditions, we came to the conclusion that the function of the air was to prevent the residual water vapour in the quartz from escaping when the bulb was heated for the purpose of observing fluorescence. The amount of water vapour capable of destroying the fluorescence is incredibly small, and in the bulb at 10^{-4} mm. pressure, sufficient comes out at 250° to accomplish this, in spite of the long preliminary baking-out process at a much higher temperature.

Niewodniczanski has recorded an observation of the mercury atomic line 4358 in the spectrum of the vapour excited by the aluminium spark, but no trace of 4046 or 5461, a very improbable circumstance, since the three lines originate from the same level. This is very important and surprising, since no similar observation has ever been recorded in spite of the very large amount of experimental work by numerous investigators. It was only after some days of experimenting that we got a faint suspicion of the line with a small quartz spectrograph, the difficulty being due to the impossibility of getting rid of the aluminium spectrum scattered by the wall of the bulb. On substituting a large spectrograph of 1 metre focus, we found to our surprise that practically all of the stronger arc lines of mercury were present, and that these increased in intensity in comparison to the intensity of the fluorescent bands, as the spark was moved closer to the bulb. Employment of the wire-gauze method used by one of us in an earlier investigation on the relation existing between the intensity of the exciting light, and the intensities of optically excited lines resulting from one-, two- and three-stage absorption processes (' Nature,' November, 1928), we found that the arc lines increased with the square of the intensity of the exciting light (aluminium spark) with the exception of the 2536 line, which increased as the first power.

---

* Vol. 121, p. 283.

† Vol. 47, p. 330.

As the spark was very close to the bulb it was necessary to make sure that no electrical discharge was excited in the bulb by induction. We assured ourselves of this by interposing a thin piece of black paper, carefully dried by heat, between the spark and the bulb. This completely destroyed all luminosity of the vapour. We also wrapped a bulb in wire gauze, connected to earth to protect it from electrical disturbances, and still obtained the arc lines. The fluorescent spectrum is shown on Plate 14, fig. I (lower spectrum), with the mercury arc lines indicated. The upper spectrum is that of the aluminium spark. Fig. II shows the continuous band which, according to Lord Rayleigh,* begins at $\lambda = 2535 \cdot 9$. On the original negative $2536 \cdot 5$ is seen distinctly reversed. The three mercury lines 2652, 2654 and 2655 are also present.

Niewodniczanski has suggested that λ 4358 was due to a two-stage absorption process. To test this we used a piece of wire gauze which reduced, non-selectively, the intensity of all light transmitted through it to one-fifth of its original value. This gauze was mounted first between the spark and the bulb, and a photograph taken of the fluorescence spectrum, the time of exposure being 3 hours. Next the gauze was put between the bulb and the slit of the spectrograph, the plate-holder lowered slightly, and a second exposure of 3 hours made. Similarly, the plate-holder was further lowered and one more exposure made of 36 minutes (one-fifth of 3 hours). If λ 4358 were due to a one-stage absorption process its intensities in the first two photographs should be equal. If due to a two-stage absorption process or to collisions between two excited entities as will be explained presently, its intensity in the second should be five times that in the first, and its intensities in the first and third should now be equal. The latter was found to be the case (see Plate 14, fig. III). In the upper photograph the gauze was between the spark and the bulb, and in the lower two it was between the bulb and the slit of the spectrograph. Fig. IV shows the more refrangible portion of the spectra. The band near 2536 is seen to have the same intensity in the two upper spectra, while the pseudo triplet at 2652 has the same intensity in the first and third. By the same method we found that the visible (green) fluorescence was due to a one-stage process. To account for the appearance of λ 4358 in the fluorescence spectrum, Niewodniczanski suggested that an electron transition from 1P to 1*s* occurred by absorption of $\lambda = 12072$, the return of the electron from 1*s* to $2p_2$ being accompanied by the emission of λ 4358. This would imply the existence of radiation of exactly $\lambda = 12072$ in the aluminium spark. To test this hypothesis he passed the light from the aluminium spark through a quartz absorption vessel containing a layer of water 11 mm. thick, which would reduce the intensity of the radiation at λ 12072 to 1/26 of its initial value. He states that "the intensity of the fluorescence excited through this water filter was but a little weaker than without filter, whereas the line λ 4358 A.U. disappeared completely." In repeating this experiment we found that the empty quartz absorption vessel alone reduced considerably the intensity of the fluorescence, and we therefore substituted a fluorite cell 11 mm. in thickness, which, when empty, did not reduce the intensity perceptibly, but when filled with water reduced greatly the intensity of the visible fluorescence. This was not in accordance with Niewodniczanski's observation quoted above, but it served to explain the non-appearance of λ 4358. For, since λ 4358 varies with the square of the intensity of the exciting light, whereas the continuous green fluorescence varies as the first power, it follows that if the water cell reduced the green fluorescence to, say, one-tenth of its original value, it would reduce the intensity of λ 4358 to one one-hundredth; in other words, no trace of it would be observed. This experiment, therefore, in no way supports Niewodniczanski's theory of the origin of λ 4358.

---

* ' Roy. Soc. Proc.,' A, vol. 111, p. 456 (1926).

It had been found previously by one of us* that when the fluorescing vapour was strongly heated, the green fluorescence was destroyed. Further, Lord Rayleigh† found that when the vapour, which was *distilling* away from a mercury arc, was passed through a red-hot tube, the continuous band in the green was destroyed, but the band with its maximum at about λ 3300 was unaffected. We investigated the effect on this latter band, which appears also in the fluorescence spectrum, of strongly heating the mercury vapour, while its pressure was kept constant. A quartz bulb was constructed with a side-tube about 4 cm. long. A drop of mercury was introduced and the bulb very thoroughly baked out and exhausted. This was suspended in the asbestos chimney with the side-tube hanging vertically down. A flame was now played over the bulb, driving all liquid mercury down into the side-tube. The temperature of this tube was that of the chimney, and this determined the vapour pressure in the bulb. The aluminium spark could now be started and an exposure made. The flame was kept playing over the bulb to maintain a temperature which would just destroy the green fluorescence. When this was done it was found that the λ 3300 band was *enormously enhanced.* A series of photographs was now taken for chimney-temperatures, varying from room temperature to 278° C. A photograph was first taken of the fluorescence spectrum when the whole bulb was at the temperature of the chimney, and immediately below this, one with the bulb superheated. The chimney-temperature was then slightly raised and another pair of photographs taken, and so on. In each case an exposure of 5 minutes was given. On Plate 14, fig. V, the lower photograph shows the spectrum with the bulb at 278°. In the upper photograph the chimney-temperature was maintained at 278°, but the bulb was superheated. The effect of superheating in suppressing the green band and enhancing the λ 3300 band is clearly shown. The band at 2536 is nearly destroyed by the superheating, as shown in the small photograph at the left, which was printed from the same plate. The photographs with the vapour *not* superheated showed that the intensity of the green band increased until a temperature of about 260° was reached. There was no noticeable difference between the intensity of 260° and that at 278°. Similarly, with the vapour superheated, the intensity of the λ 3300 band increased up to a chimney-temperature of about 260°. There was again no noticeable difference in intensity for chimney-temperatures of 260° and 278°.

---

* Wood, 'Phil. Mag.,' vol. 18, p. 247 (1909).

† 'Roy. Soc. Proc.,' A, vol. 114, p. 633 (1927).

This series of photographs showed also that $\lambda = 2536$ appeared first at a temperature of 100°. Lord Rayleigh, though expressing the opinion that the 2536 line was present in the spectrum, drew attention to the circumstance that the head of the band at 2535·9 might easily be mistaken for it, especially if a wide slit were employed. He regarded the emission at this region as consisting of two parts, a band shaded off towards the red with its head at 2535·9 and a line at 2540, which should perhaps be termed a band head. One of his photographs shows an emission band, with dark reversals at both of these wave-lengths (Plate 8, fig. 2, p. 460, 'Roy. Soc. Proc.,' vol. 111). This led him to place the head of the continuous emission band at 2535·9. Our results seem to show also that the atomic line 2536 really appears in the emission spectrum, and that the radiation at 2535·9 in Lord Rayleigh's photograph is the short wave-length edge or wing of the broadened 2536 line, its centre removed by absorption. Our reasons for this conclusion are as follows:—

At low vapour densities we obtain 2536 as a "hair fine" line in coincidence with and of a narrowness equal to that of the emission line of a water-cooled mercury arc. As a matter of fact, the 2536 line must appear in the emission spectrum if the other arc lines of the element are present as in this case. At higher densities of vapour a continuous band develops on the long wave-length side of 2536, which, however, does not shade off gradually, as is usual with band spectra, but is of nearly uniform density, and terminated sharply at the end away from 2536, so that it has the appearance of a line photographed with a wide open slit. This condition is shown on Plate 14, fig. VI. Superposed on this there is also a band shading off towards the red which extends to some distance beyond the band of uniform intensity. This is evidently the emission band corresponding to the absorption band shown by dense mercury vapour. The same condition is shown also in one of Lord Rayleigh's photographs (figs. 3 and 4, *loc. cit.*).

The most probable explanation is that we have a band heading on 2539 shading off towards 2536 and a band heading on or very near 2536 and shading off towards and beyond 2539. The superposition of these two bands will exactly duplicate the appearance shown by the photographs, including the region of minimum intensity at the centre of the band, *i.e.*, midway between 2536 and 2539. A diagram of the two intensity curves for fluorescent emission is given in fig. 1*a*.

We have photographed the absorption spectrum of rather dense mercury

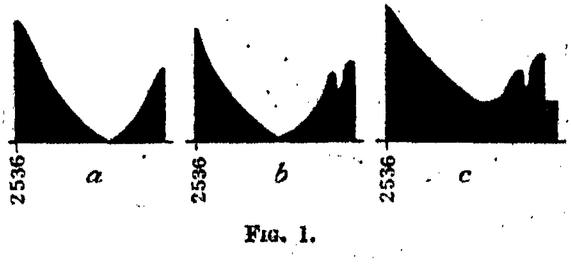
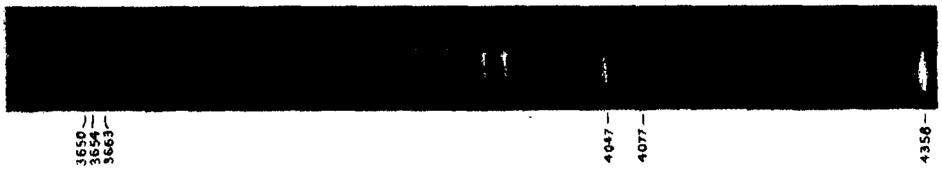
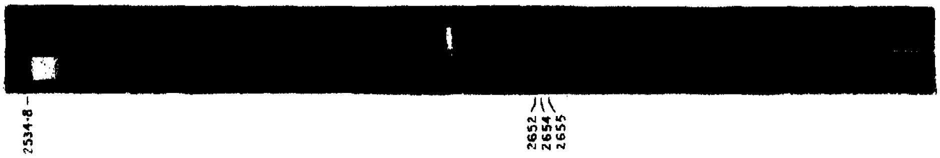


Fig. 1.

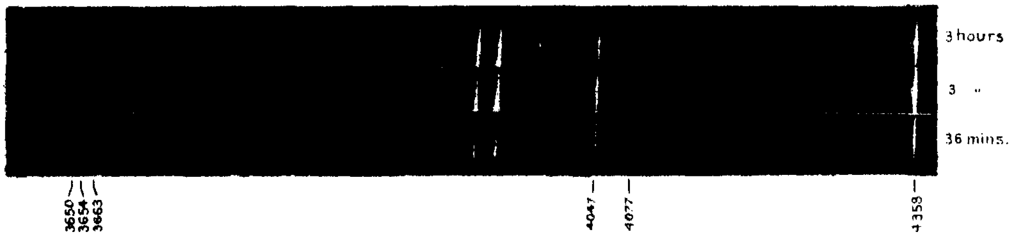
I



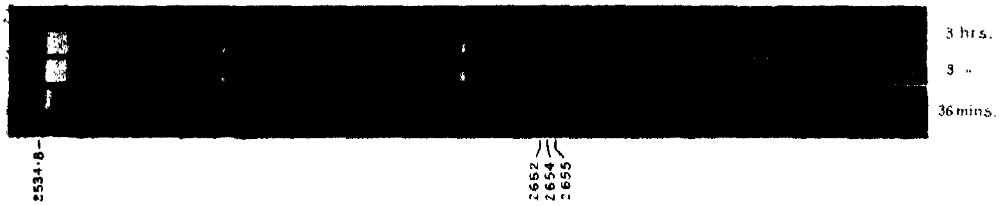
II



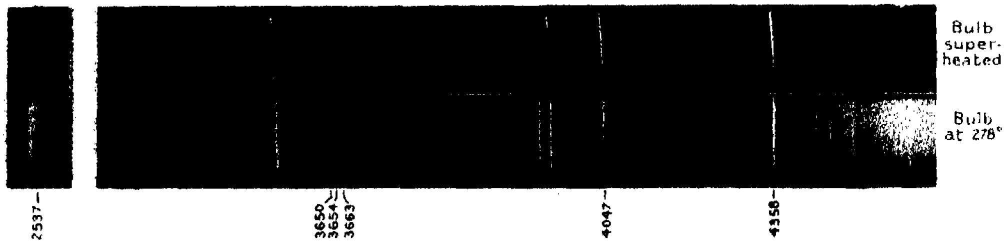
III



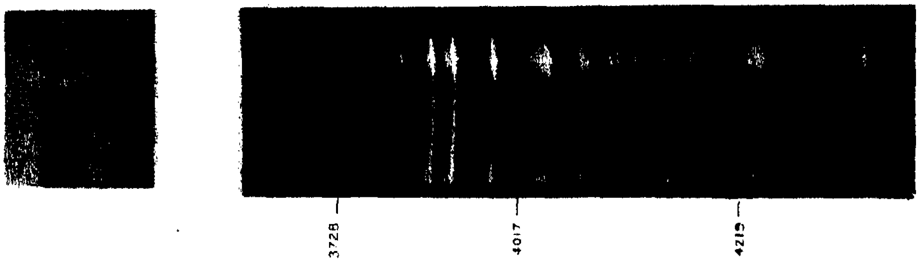
IV



V



VI



vapour, and the 2539 band clearly shades off towards the short wave-lengths, but not quite uniformly, there being a distinct indication of a narrow region of less powerful absorption close to the head of the band, the two bands appearing as in fig. 1*a*. At a slightly higher density a narrow region of uniform absorption appeared on the long wave-length side of the 2539 band, as shown in fig. 1*b*, and at still higher densities the band headed on 2536 has broadened to such a degree that all trace of these smaller structures is wiped out, fig. 1*c*.

The study of the effect of superheating on the 2536 bands gave rather contradictory results. On Plate 14, fig. V, the superheating has almost abolished the bands, and on some plates we found that superheating destroyed the 2536 lines entirely, but left a narrow continuous band of uniform intensity. On trying to verify these results with some new bulbs, we were unable to find any effect at all of superheating, and suspect that the earlier results may have involved the presence of impurities. The observations should be repeated with a double oven, one for controlling the density of the mercury vapour and the other for the superheating. The tube should also be in communication with the pump, so that any water vapour liberated can be removed. Until this has been done it is useless to try to interpret the results.

The arc lines are all increased in intensity by superheating (with the exception of 2536 for densities of mercury vapour corresponding to temperatures below 250°) but for higher densities (temperatures above 265°) the lines 4046 and 4358 are decreased by superheating, while the others (which originate from the fall of electrons from higher levels) continue to be increased.

The mercury hydride bands, which appear when water vapour is present, behave in the same way as the lines 4046 and 4358, *i.e.*, superheating increases them for densities below 250° and decreases them for temperatures above 265°. These bands are shown on Plate 14, fig. V. With more water vapour in the bulb the hydride bands are more strongly developed, as shown by Plate 14, fig. VI, the upper spectrum being that of the aluminium spark. The hydride band heads at wave-lengths 3728, 4017 and 4219 are indicated on the photograph. The arc lines (2536 excepted) and the mercury hydride bands and HO bands alone increased with the square of the intensity of the exciting light. It does not seem probable that a two-stage absorption process is responsible for this. Lines or bands which result from the collision of two excited entities will behave in the same way, for if we double the intensity of the exciting light, we double the number of excited entities (atoms or molecules). The collision frequency of a single one will be doubled by this process, and there are twice as many to collide; therefore, the intensity of the radiation will increase fourfold.

We found the carbon line 2478 quite strongly recorded on plate, due probably to the presence of traces of organic vapours.

An Improved Method of Ultra-Violet Polarimetry. Anomalous Rotary Dispersion of Sodium Tartrate.

By Thomas Martin Lowry, F.R.S., and Maurice Arthur Vernon.

(Received June 13, 1928.)

A large number of ultra-violet polarimetric measurements have been made by a method described by one of us in 1908.* In this method a triple-field polarimeter, with Foucault in place of Nicol prisms, is arranged in series with a quartz or quartz-calcite spectrograph. A quartz-calcite lens, replacing the eyepiece of the polarimeter, casts a real image of the triple field on the slit of the spectrograph, and thus gives rise to a triply divided spectrum, intersected by dark bands. The wave-length corresponding with a given rotation is determined by finding a line which is of equal intensity in the three fields. The approximate form of the curve of rotatory dispersion can be determined by setting the analyser in a series of positions separated by 5° or 10°; but it is then advisable to "bracket" some of the more conspicuous lines by making fresh exposures at intervals of perhaps 0·2°, when the rotation corresponding with the wave-length of a given line can be determined within about 0·1°. This degree of accuracy is less than that which can be reached in the central part of the visible spectrum, where the readings may be reproduced under favourable conditions within about 0·01°; but it is not appreciably less than the accuracy of visual readings in the red and violet regions, and for many purposes is quite satisfactory. Thus, with a column of quartz 496 mm. in length, readings to 1° sufficed to give the rotatory power in degrees per millimetre within 0·002°, corresponding with an error of about 1 part in 100,000 at wave-length 2327.†

---

* Lowry, 'Roy. Soc. Proc.,' A, vol. 81, p. 472.

† Lowry and Coode-Adams, 'Phil. Trans.,' A, vol. 226, p. 391 (1927).

The present investigation was undertaken in order to find out whether the same apparatus could be used to record with sufficient accuracy the much smaller rotations of solutions which had been largely diluted, in order to render them transparent in a region near to (or covered by) an absorption band. For this purpose the concentration is often reduced to 1 per cent.: the observed rotations may then be of the order of 0·5°, and must be read to 0·01° or less in order to give a true impression of the form of the curve of rotatory dispersion.

The method adopted was to take a series of photographs for settings of the analyser at intervals of about 0·1°, in the same way as before, but to measure the optical density of the photographic images by means of a densitometer, instead of relying only on a visual comparison. The theoretical values of the ratio of the intensities of the outer and central portions of the triple field, for settings at intervals of 0·1° with a half-shadow of 5°, are shown in fig. 1, A. If the density of the photographic image were a linear function of the intensity of the light, it would be a simple matter to deduce the exact position of the "extinction" from widely spaced readings on either side; but, in view of the peculiar form of the Hurter-Driffield curves, this linear relationship does not hold even approximately. Since, therefore, the different spectral lines vary widely in intensity, the position of the extinction must be deduced exclusively from readings in which the ratio of the densities of the different parts of the field does not differ much from unity, when the error from this cause disappears completely.

Trials were made with a mercury arc as a source of light, but the iron arc, which had been used in all the earlier measurements, was found to provide more practicable conditions. In view of the small size of the photographic image of the triply divided lines, it was convenient to measure the photographic densities on an enlargement, and to work with a wider slit than when only a visual estimate was made of the relative intensities of the different parts of a line. The densities were measured on a B.P.R.A. densitometer.[*] The lag of the selenium cell made the measurements rather tedious, but the results were quite trustworthy, and could be reproduced very accurately. A typical series of measurements, selected at random, is given below, the settings being separated by 0·1°.

Setting	θ	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
Density-ratio	ρ	1·514	1·316	1·257	1·185	1·072	1·031	0·994	0·946	0·897

---

[*] Toy and Rawling, 'J. Scient. Inst.,' vol. 1, p. 362 (1924).

These numbers are shown graphically in fig. 1, B, where the relative uniformity of the density-ratio suggests that the lines examined were on the "over-exposure" part of the Hurter-Driffield curve. The variable intensities of the lines, and the irregularity of illumination in successive exposures, make it impossible to plot a smooth curve between the setting and the density-ratio; but by using only those ratios which are nearest to equality, namely, between 0·9 and 1·1, the position of the extinction can be read off to 0·005°. The ultra-violet readings are then at least as accurate as those obtained by visual observations even in the central part of the visible spectrum.

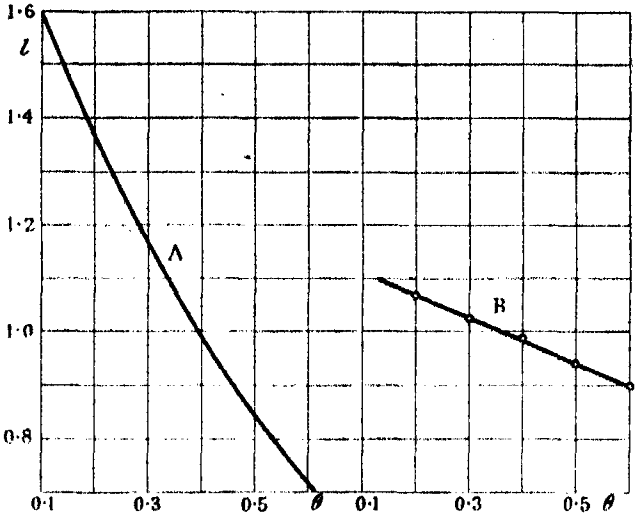


Fig. 1.—Curve A, calculated Ratio of Intensities of Light. Curve B, observed Ratio of Photographic Densities.

This method of measurement was used to study the character of the rotatory dispersion of sodium tartrate in the ultra-violet region. This salt very nearly obeys Biot's Law,* but small systematic deviations from the law of inverse squares indicate that the rotatory dispersion is probably complex. An equation was therefore deduced of a type which would give rise to anomalous dispersion in the region of shorter wave-lengths. The correctness of this prevision is shown by the data given in Table I, and plotted in fig. 2, for a solution of about 1 per cent. concentration. This result differs from those of Nutting,† who observed a normal rotatory dispersion in Rochelle salt up to 3090 Å.U., but it is in agreement with the recent extensive observations of Descamps ('Thesis,' Brussels, 1928).

---

* Lowry and Austin, 'Phil. Trans.,' A, vol. 222, p. 249 (1922).

† 'Phys. Rev.,' vol. 17, p. 1 (1903).

Table I.—Rotatory Dispersion of Sodium Tartrate (0·00935 N.) at 20°. Concentration 0·963 gramme per 100 grammes of solution. Density $d^{20°}_{4°} = 1·0044$. Length of tube = 1 dm. Half-shadow angle about 5°.

λ.	α.	λ.	α.
Å	°	Å	°
4919	0·20	3684	0·53
4384	0·39	3586	0·55
4272	0·38	3441	0·52
4144	0·45	3306	0·12
4045	0·46	3213	− 0·25
3926	0·50	3100	− 0·77
3786	0·51	2983	− 1·50

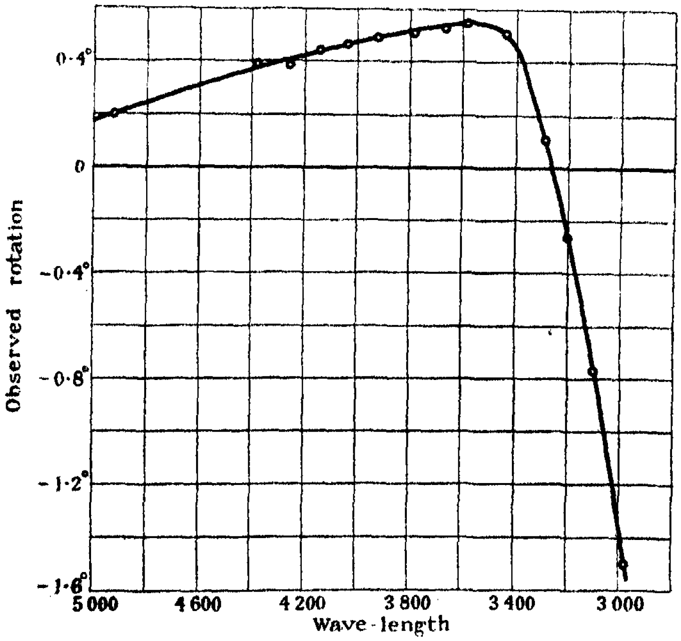


Fig. 2.—Observed Rotations of Sodium Tartrate.

ERRATA.

Paper by F. H. Constable.

P. 197, line 13.	*For* heterogeneity *read* homogeneity.
P. 199, line 8.	*For* $\frac{p(2n-1)\lambda}{4\mu}$ *read* $\frac{4\mu}{p(2n-1)\lambda}$.
,, ,,	*For* $\frac{pn\lambda}{2\mu}$ *read* $\frac{2\mu}{pn\lambda}$.

Paper by T. M. Lowry and G. Owen.

P. 505, Title.	*For* G. Glyn Owen *read* Glyn Owen.
P. 515, lines 13, 14.	*For* 0·031, 0·122 *read* 1·031, 1·122.
P. 515, line 37.	*For* 75 days *read* 39 days.
P. 515, line 39.	*For* 0·000006 *read* 0·000016.
,, ,,	*For* four *read* somewhat.
,, ,,	*For* 250 *read* 100.
P. 516, lines 5, 6.	*Omit* which was . . . tube.
P. 516, Table IV (*a*).	*For* 0·0000065 *read* 0·0000165.
,, ,,	*For* 0·0000064 *read* 0·0000164.
P. 516, Table IV (*b*).	*Against* 935 mins. *read* — rotation.
,, ,,	*Drop all other rotation figures one line lower.*

OBITUARY NOTICES.

CONTENTS.

PAGE

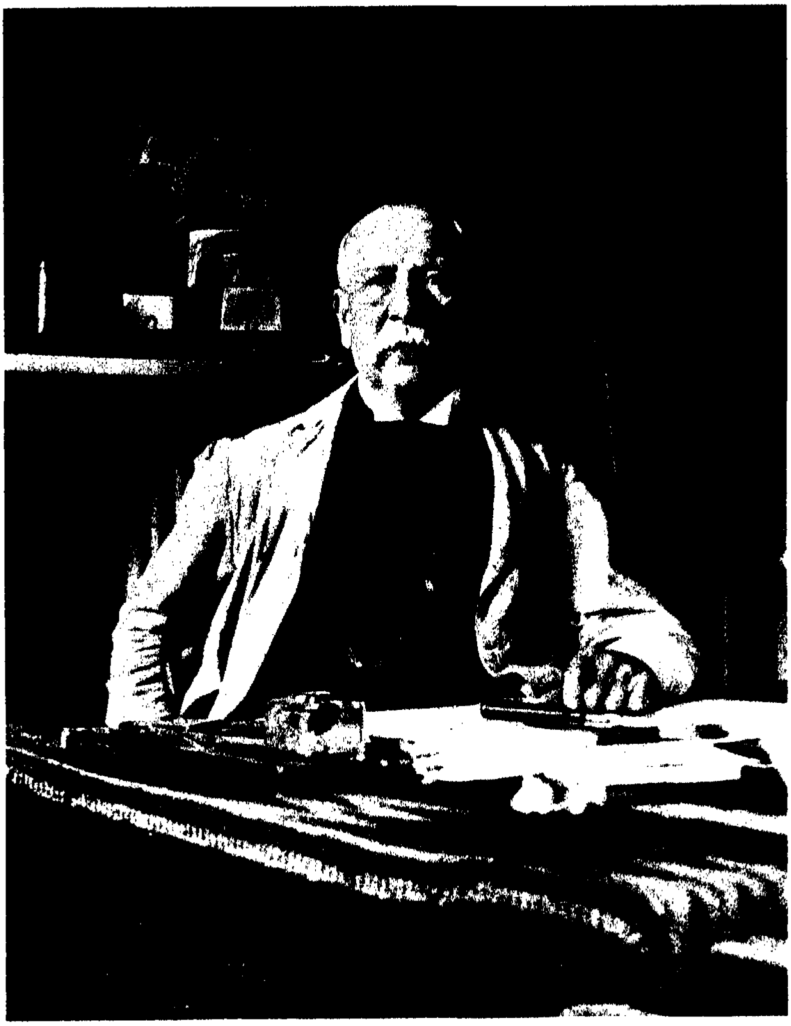
Alfred George Greenhill (with portrait) i

Gösta Mittag-Leffler (with portrait) v

Svante August Arrhenius (with portrait) ix

Paul Heinrich Ritter von Groth (with portrait) xx

William Henry Dines (with portrait) xxiii



A. G. Greenhill

ALFRED GEORGE GREENHILL, 1847–1927.

Alfred George Greenhill, born November 29, 1847, elected F.R.S. in 1888, died February 10, 1927. After a distinguished career at Christ's Hospital, and at St. John's College, Cambridge, he was second wrangler in 1870, but bracketed with the Senior Wrangler, Richard Pendlebury, of his own college, in the Smith Prize Examination. In the same year he was elected Fellow of the College. Shortly afterwards he was appointed to the Royal Indian Engineering College, Cooper's Hill. In 1873, however, he was made Fellow and Lecturer at Emmanuel College, Cambridge. This he left in 1876 to become Professor of Mathematics to the Advanced Class of Artillery Officers at Woolwich. After more than thirty years' government service he retired, and the rest of his life was spent, as a bachelor, in London, first in New Inn, and later in Staple Inn; in his later years, he was the recipient of a Civil List Pension. During his residence in London he was very active in mathematical and scientific circles; he became a Fellow of the Royal Society in 1888, and received a Royal Medal from that Society in 1906; he was awarded the De Morgan Medal by the London Mathematical Society in 1902; he had served as President of this Society 1890–92. He was on the Councils of both societies, of the Royal Society, in 1896, 1897, and of the London Mathematical Society for many years. He also became a corresponding member of the Académie des Sciences (Paris), and a foreign member of the Accademia dei Lincei (Rome). He was knighted in 1908, on his retirement from Woolwich.

His original contributions to knowledge were mainly to Dynamics, to Hydrodynamics, and to Elasticity. Of the great value of these a highly appreciative account is given by an expert, in Prof. Love's Obituary Notice of Greenhill, "Journal of the London Math. Soc.," vol. iii, pp. 29 and 30, 1928. Prof. Love writes as follows:—

"Greenhill made valuable contributions to hydrodynamics. In the theory of the motion of a solid in a fluid, developed by Kelvin, Kirchhoff, and others, the fluid is regarded as acting only by pressure and inertia, and obeying the electrical law of flow. Greenhill applied this theory to give an account of the steadiness of flight conferred upon an elongated projectile by rifling. He determined the least angular velocity about its axis for which steady motion of a solid of revolution, moving in the direction of its axis, can be stable. This was done incidentally in a paper on 'Fluid motion between confocal ellipsoids and confocal elliptic cylinders' ('Quart J. of Math.,' vol. 16 (1879), p. 227). The ballistic implications of the result, in regard to the degree of rifling required for various types of projectiles, were pointed out in his article 'Hydromechanics' in the 10th edition of the 'Encyclopædia Britannica,' where it is stated that they agree fairly well with what has been found in practice. Probably there was some earlier publication of these results. This practical application of what was regarded as a recondite mathematical theory earned for him much renown at Woolwich. Some years later he took up the theory which traces fluid resistance to discontinuity, and wrote a long 'Report on stream line motion past a plane barrier' (1910) for the Advisory Committee on Aeronautics. In this he gave an account of the theory, as previously developed, and extended it by completing the solutions of various special problems that require elliptic functions. At that time progress was just beginning to be made with the difficult problem of curved boundaries, and he wrote afterwards a supplementary report (1916), dealing with the researches in this department up to date, and completing many of them in the same way. But his contributions to hydrodynamics do not begin and end with applications of elliptic functions. The rifling problem does not involve them, though the paper in which it was solved introduces them, and they play no part in the investigation that he made of the problem, first attacked by Lejeune Dirichlet, of the motion of a liquid ellipsoid, subject to its own attraction, and moving so as to remain ellipsoidal. Riemann had shown that in all such motions the co-ordinates of a fluid particle must be linear functions of their initial values, and had obtained the equations which the coefficients of these linear functions must satisfy. Greenhill showed that all the possible motions in which the ellipsoid retains its form can be obtained by compounding irrotational motion of liquid in a rotating ellipsoidal shell with vortex motion of the liquid rotating as if rigid, thus interpreting Riemann's result for such motions, and arriving at it by a more intuitive method ('Proc. Camb. Phil. Soc.,' vol. 4 (1883), p. 4).

"Perhaps the most striking of Greenhill's contributions to the theory of elasticity is his paper on '. . . . height consistent with stability . . . . ' ('Proc. Camb. Phil. Soc.,' vol. 4 (1883), p. 65). Here he attacked for the first time the question of the greatest length which a shaft standing upright can have if it is not to be bent by its own weight. In Euler's well-known strut problem the weight of the shaft is neglected, and the tendency to bend arises from a supported load. The problem turned out to be easy, and Greenhill made applications of the result to the question of the greatest height to which a tree can grow. As always, he took great pains to ascertain the observed facts, and, in this case, he collected interesting data in regard to the heights of exceptional trees, comparing them with his result. Another important contribution to the theory of elastic stability was his investigation ('Proc. Lond. Inst. Mech. Engineers' (1883), p. 182) of the condition for the whirling of shafts, in regard to which he brought out the importance of the gravest period of lateral vibration. He made out that a long shaft, rotating between bearings, can remain straight if its rotation period exceeds this gravest vibrational period. If it rotates too fast it may whirl. Reference has already been made to his elaborate discussion of the problem of the deformation of an elastic circular ring by pressure ('Math. Ann.,' vol. 52 (1899), p. 465). At the time it was hoped that the solution of this rather ideal problem would throw light on the baffling, but technically important, problem of the collapse of boiler flues, but the solution of this was to be found much later, along quite different lines, and by another hand."

But Greenhill's contributions to the mathematical and technical journals were many and voluminous. A large part of many of his papers is taken up with detailed computations with elliptic functions. One can respect his desire to present the solution of his problems in finished form, and can understand his enthusiasm when one reflects that he was largely a pioneer in making known the applications of these functions; but work of this kind is apt to miss appreciation on both sides. To the physicist it means the necessity of understanding much detailed algebra, and is less convincing than an approximate solution by elementary functions; to the mathematician its special character and lack of horizon, all that makes it suitable finally for numerical estimates, renders it extremely cramping and dull. Moreover, in Greenhill's time there were no adequate numerical tables of the elliptic functions; it is understood that it is due largely to his efforts that the Smithsonian Institution undertook to supply this want, and so placed the scientific world under great obligation.

To judge of his mathematical achievements one must in fairness recall the outlook of the English Universities in mathematics at a time, say, about forty years ago. It is clear that to his contemporaries of that time he was an advanced thinker. But with tastes which were apparently primarily mathematical he was impatient of systematic theory, and desired always to be in touch with some practical application, without being willing to give himself to prolonged experimental work. To many in the practical world he stood, it would seem, as the apostle of higher mathematical methods; to the mathematicians in England, however, whose theoretical outlook was constantly being enlarged, he was persistent in his denunciation of much that had emerged since his own student years, unless capable of deduction by the algebraic methods in which he was so great an adept.

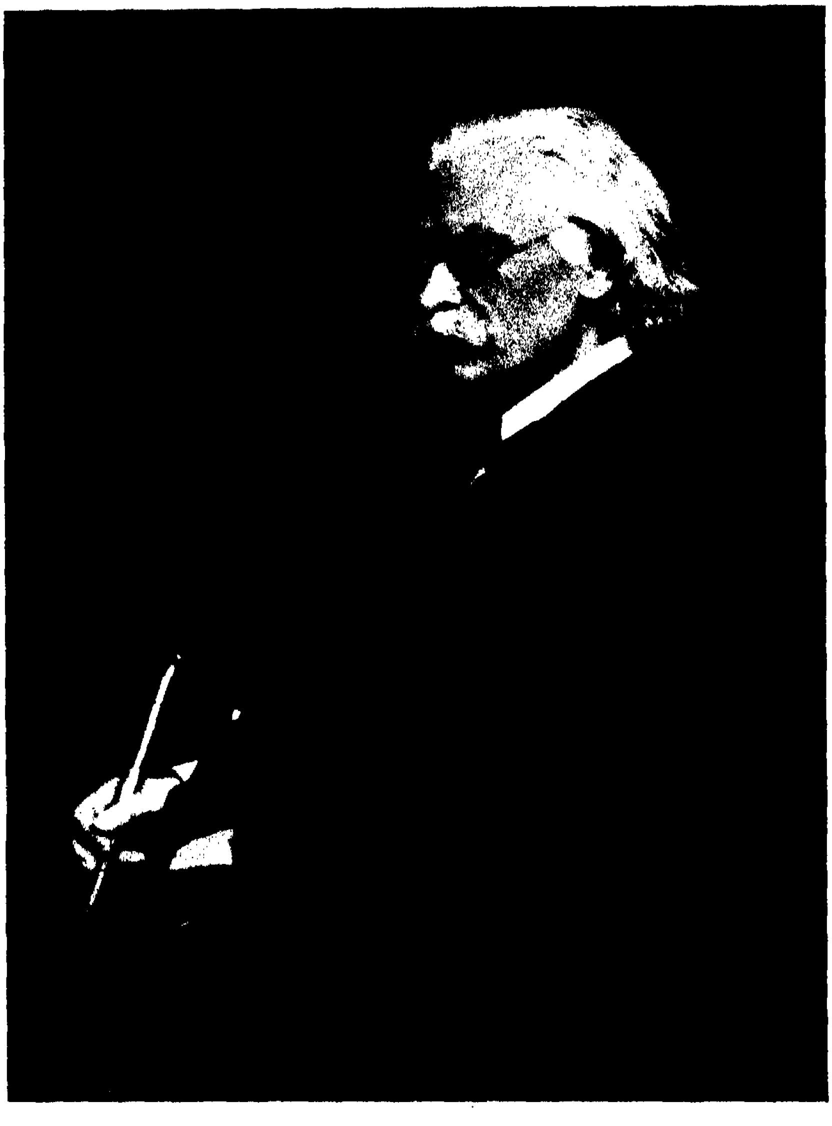
His characteristics are indelibly printed on the text books he wrote, all of them arresting, and full of points of novelty—books that have a permanent value as milestones in our pædagogy of mathematics in England, but each requiring special study of its method of arrangement. His second, on the applications of Elliptic Functions, 1892, which was translated into French, was a pioneer work: it develops the addition theorem in connection with the motion of a pendulum, and explains Weierstrass' factors for the sigma function in connection with the plane flow of electricity. This is very interesting to anyone, and seems admirable from the point of view from which the book was written;

but it is worse than inadequate in its mutilation of a wonderful theory, painfully won over many years by many workers, by more germane methods. His first book, on the 'Differential and Integral Calculus,' 1886, is smaller, but also practical, especially in its introduction of integration concurrently with differentiation. His most valuable book, probably, on 'Hydrostatics,' 1894, is enriched by the introduction of many of the actual problems of the science, and marked a great advance on the material which had been called by the same name for many years. All these books have most carefully drawn diagrams, costing the author infinite trouble.

But, apart from Greenhill's scientific discoveries, and the influence of his educational outlook, there is more to say, and that the best which can be said of any man. His personality, apart from a certain wilful pose of eccentricity (*cf.* 'The Times Obituary,' 14 February, 1927), was, especially to younger men, and in the earlier years of his professorial activities, inspiring by its earnestness in a manner not easy to put into words; his kindliness, and interest in the work of others, made a great impression on his friends; he will be greatly missed.

H. F. B.

The portrait, here reproduced, of Sir George Greenhill was taken in August, 1916, in his study at Staple Inn, by Mr. J. W. Hicks, M.B.E., F.R.A.S., who kindly lent the photograph.



Mittag-Leffler

GÖSTA MITTAG-LEFFLER, 1846–1927.

Magnus Gösta Mittag-Leffler, who died last year at the age of eighty-one, the oldest of our foreign members, was for many years the recognised leader of the Scandinavian school of mathematics.

Mittag-Leffler was a remarkable man in many ways. He was a mathematician of the front rank, whose contributions to analysis had become classical, and had played a great part in the inspiration of later research; he was a man of strong personality, fired by an intense devotion to his chosen study, and he had the persistence, the position and the means to make his enthusiasms count. He thus won for himself a position unlike that of any other Scandinavian mathematician, and one which it was a little difficult for anyone who had not met him in his own country to realise. I can remember well the occasion when he lectured for the last time to a Scandinavian Congress, at Copenhagen in 1925, and the whole audience rose and stood as he entered the room. It was a reception rather astonishing at first to a visitor from a less ceremonious country; but it was an entirely spontaneous expression of the universal feeling that to him, more than to any other single man, the great advance in the status of Scandinavian mathematics during the last fifty years was due.

Mittag-Leffler had been overwhelmed with honours; he was a doctor, for example, of four English or Scottish universities; he had been a president or vice-president of endless congresses; and he was an honorary member of almost every scientific society in the world.* He travelled continually, and there was no mathematician who was so familiar a figure abroad. But to the outside world he was, above everything, the editor of the 'Acta Mathematica,'† the famous journal which he founded, and, with the co-operation of a committee of the four Scandinavian countries,‡ edited for forty-five years.

Periodicals are the most important material facts in the mathematical world. The 'Acta' took its place in the front rank of mathematical periodicals from the beginning, and its rank and standard have never declined. It has always been the most completely international of all mathematical journals. All such journals are international to some extent; 'Crelle's Journal' made its reputation on the works of Abel, and the leading German periodicals of the present day, the 'Mathematische Annalen' and the 'Mathematische Zeitschrift,' are very largely international. The 'Acta' was always international in a more thorough-going sense than this. It would hardly have been practicable, in 1881, to support a mathematical journal of the first rank on Scandinavian mathematics alone; and Mittag-Leffler understood this, and set himself from the beginning a standard which, without the co-operation of mathematicians of all countries, it would have been quite impossible to maintain.

---

* He was a foreign member of the Royal Society from 1896 until his death.

† A fuller account of Mittag-Leffler's activities in connection with the 'Acta' will be found in Nörlund's notice of him in vol. 50. Prof. Nörlund very kindly allowed me to see the proof-sheets of this notice, and I have drawn from it freely. The portrait here reproduced is also copied from the same notice by Prof. Nörlund's permission.

‡ Including Finland.

Two great mathematicians in particular, Poincaré and Cantor, contributed freely to the early volumes. It is difficult to realise to-day how Cantor had to struggle for recognition. The mathematicians of his time, however, were singularly slow to understand the overwhelming importance of his work; and Mittag-Leffler, a younger man, was one of the first to recognise it, and to apply Cantor's ideas successfully in the region of "ordinary" analysis. In Mittag-Leffler's last published paper he recalls this with very justifiable pride.* But Mittag-Leffler was always a good judge of the quality of the work submitted to him for publication. Even in his later years, when most of the editorial work was delegated to others, he retained that curious sense which enables the great editor to feel the value of work at which he has hardly glanced, and it was this gift which enabled him to carry the heavy responsibility which he had assumed in a way which would hardly have been possible to any other mathematician. It was a heavy responsibility in more ways than one, for the finances of the 'Acta,' like those of other journals, were imperilled by the war, and there was a period when Mittag-Leffler had to supplement the Government subventions heavily himself. Like other journals, the 'Acta' has overcome its troubles, and Mittag-Leffler could no doubt feel, when the last three volumes were dedicated to him as a *Festschrift* on his eightieth birthday, that the future of his great foundation was secure.

Mittag-Leffler's own work is occupied almost entirely with the general theory of functions analytic in Weierstrass's sense. He was an excellent writer. His account of the foundations of the theory of elliptic functions, modelled, no doubt, on Weierstrass's lectures which he had attended in Berlin two years before, published first in Helsingfors in 1876, and translated into English in 1923 in the 'Annals of Mathematics,' is a really admirable piece of exposition, and makes one regret that he never wrote a treatise on a larger scale. He occupied himself on various occasions with fundamentals in the theory of functions. For example, he gave a proof of Cauchy's theorem in 1875, which marked at the time a very definite advance. Like Goursat after him, he divides the contour into elementary contours by a network, and argues directly from the definition of the derivative in every mesh; but he assumes what amounts substantially to uniform differentiability, and so misses what is really characteristic in Goursat's proof. He also wrote a number of memoirs on linear differential equations.

---

* A short note on a very interesting historical article by Schönflies ["Die Krisis in Cantor's mathematischem Schaffen," 'Acta Math.,' vol. 50, pp. 1-23 (1927)]. This article reproduces a number of Cantor's letters to Mittag-Leffler.

His best-known and most characteristic work, however, is concerned with the problem of the analytic representation of a one-valued function (or one-valued branch of a function) with assigned singularities, and falls into two halves associated with the periods of his life following on the years 1877 and 1900. It was in 1877 that Mittag-Leffler first published the famous "Mittag-Leffler's theorem" of the textbooks, a theorem suggested, of course, by Weierstrass's factor-theorem, but in no sense a corollary of it. In a later memoir—his first contribution to the 'Acta,' in 1884—Mittag-Leffler gave various generalisations of his theorem, of which treatises such as those of Forsyth and Osgood give a full account. In the second series of memoirs, the first of which appeared in 1900, we find the problem of analytic continuation envisaged in an entirely modern form. Mittag-Leffler considers the general problem of continuation for an element of any analytic function regular, say, at $z = a$. He introduces for the first time the idea, now well established as of fundamental importance in all such investigation, of the "star" (*Hauptstern*) associated with a; the region obtained by removing from the plane the further segments of all radii issuing from a and passing through singular points. He solves in various manners and in explicit form the problem of representing the function by an analytical expression convergent throughout the whole interior of the star. Some of these solutions are series of polynomials in z, while others are limits of associated functions, such as

$$\lim_{\alpha \to 0} \sum_{n=0}^{\infty} \frac{c_n z^n}{\Gamma(1 + \alpha n)},$$

where we suppose that a is zero and that $\Sigma c_n z^n$ is the Taylor's series of $f(z)$. Representations similar to the last, but involving slightly different convergence factors, such as $n^{-\alpha n}$, have also been used with much effect by other writers, such as Lindelöf and Le Roy. A glance at Bieberbach's article in the 'Enzyklopädie der Mathematische Wissenschaften' (II C 4) is enough to show the range of later investigations which have their root in Mittag-Leffler's researches.

Mittag-Leffler had two homes between which he divided his time in Sweden. His principal residence was in Djursholm, in the suburbs of Stockholm. Here was his library, the finest mathematical library in the world, perhaps—at any rate for a working mathematician—and now in the hands of the Mittag-Leffler Institut and the Swedish Academy of Sciences. All books and periodicals were there (or all, at any rate, that any sane man could want), and a colossal collection of "separata"; and if one got tired one could read the correspondence of all the mathematicians in the world, or enjoy the view of Stockholm from the roof. He had also a country house at Tällberg, some two hundred miles to the north, in the beautiful and rather primitive Dalarne country. It was a delightful experience to stay there, and there perhaps Mittag-Leffler appeared at his best, a most entertaining mixture of the great international mathematician and the rather naive country squire. He was a strong nationalist, in spite of his internationalism, as anyone who lived in so beautiful a country well might be; and he loved his house and his garden, and his position as the landowner of the countryside. Even his tennis court, " the most northerly tennis court in the world "—that shapeless mass of disintegrating rubble, with holes in which the ball remained stationary where it pitched—he regarded with an entirely disarming pride. I am sure there can be no one who has ever stayed at Tällberg who is not saddened by the thought that he will never stay there again.

Mittag-Leffler played a very great part in the history of the mathematics of his time. He had exceptional opportunities, and exactly the blend of qualities required to take advantage of them to the full. There have been greater mathematicians during the last fifty years, but no one who has done in his way more for mathematics.

G. H. H.

Svante Arrhenius.

(By permission from the Journal of the Chemical Society.)

SVANTE AUGUST ARRHENIUS, 1859–1927.

In his address to the British Association in 1884 the President, Lord Rayleigh, said: "From the further study of electrolysis we may expect to gain improved views as to the nature of chemical reactions, and of the forces concerned in bringing them about. . . . I cannot help thinking that the next great advance, of which we have already seen some foreshadowing, will come on this side."

The first step in the advance spoken of by Lord Rayleigh had already been made, for in that same year the young Swede, Arrhenius, published in the *Bihang* of the Swedish Academy of Sciences a memoir with the title "Recherches sur la conductibilité galvanique des électrolytes. Première partie: La conductibilité des solutions aqueuses extrêmement diluées. Seconde partie: Théorie chimique des électrolytes." This memoir contains the germ of the theory of electrolytic dissociation, which however only received its definite statement three years later.

Svante Arrhenius came of Swedish farmer folk, a remote ancestor being one Lasse Olofsson, who in 1620 moved to the village of Årena, from which the family derived its surname in latinised form. His father, Svante Gustav Arrhenius (1813–88), studied at the University of Upsala, and later established himself in that town as a land surveyor. He was appointed collector to the University, but the emoluments of the post were to begin with so meagre that he was forced to undertake in addition the management of the estate of Wijk, on Lake Mälar. He married in 1855 Caroline Thunberg of Klövdala, and at Wijk there was born on February 19, 1859, a son whom they called Svante August Arrhenius. Owing to improved prospects the family moved to Upsala in the beginning of 1860. Young Svante was educated at the Cathedral School of Upsala, which he left at the age of seventeen, with a good record in mathematics and physics, to enter the University of Upsala, where he soon passed the candidate's examination, admitting to study for the doctorate. It seems to have been his original intention to take chemistry as his main subject under Cleve, well known for his investigations on the rare earths and on complex ammoniacal compounds; but in 1881 he definitely turned to physics, although the conditions for its study in Upsala were far from ideal. Thalén, who was at that time Professor of Physics there, did not encourage independent research work in his laboratory, and Arrhenius was forced to look about for some other opportunity to begin physical investigation. He, with a fellow student, repaired in September 1881, to Stockholm with the intention of working in the laboratory of Erik Edlund, Professor of Physics to the Swedish Academy. Edlund gave them a hearty welcome, and they began by assisting him in his work on electromotive forces in the spark discharge. In the spring of the following year Arrhenius started his first independent research on the decay of galvanic polarisation with time, an account of which was published in the *Bihang* of the Swedish Academy in 1883. From this he passed to the measurement of the conducting power of electrolytic solutions. Curiously enough, what led him to the investigation of electrolytic solutions was not directly concerned with the conducting substances themselves. He tells us that Cleve in his lectures had emphasised the impossibility of ascertaining the molecular weights of substances, such as sugar, which could not be volatilised without decomposition. Arrhenius rightly recognised that this was a great drawback, by the removal of which a considerable advance in chemistry would be rendered possible. He was unaware of Raoult's work on the same subject, and thought that some light might be thrown on the molecular weight of dissolved substances by measurements of electrolytic conductivity. He knew that when some of the water of a conducting solution was replaced by more complex non-conducting substances, such as alcohol, the conductivity was lowered, and he thought it might be feasible to deduce the molecular weight of this added substance from its effect on the conductivity. He had not proceeded far with his measurements, however, when he recognised that the state of the conducting salt was the matter of first importance.

Arrhenius completed his experimental work in the spring of 1883. He measured the resistance of a considerable number of salts, acids and bases in dilute solutions, and found for salts, as Kohlrausch had found earlier, that the specific conductivity is nearly proportional to concentration. Departure from this ideal behaviour he took as a basis for classification of the dissolved electrolytes, and showed that chemically similar substances fell into the same category when classified on this basis. A discussion of the data led Arrhenius to the conclusion that "if on dilution of a solution the conductivity does not change proportionally to the amount of electrolyte, then a chemical change has occurred on addition of the solvent." He exemplifies this by the consideration of potassium cyanide, the abnormal behaviour of which he attributes to the partial splitting of the salt into acid and base on dilution with water. The abnormal values obtained for soluble hydroxides and dilute solutions of acids he ascribes to the presence of small quantities of ammonium carbonate in the solvent water.

The importance of this paper, however, does not lie in the experimental measurements or in the detailed deductions, but in the general ideas which the author developed in the second part. He bases his theoretical treatment on the hypothesis of Williamson and Clausius, and derives from the hypothesis the notion of closed circular currents in the electrolytic solution in its normal state (*i.e.*, when not undergoing electrolysis), which are due to the momentary separation of the ions and their re-combination with other than their original partners. This notion he uses in dealing with the equilibrium between electrolytes in aqueous solutions. But by far the most important original idea, on which he founds his further treatment, is that of the distinction of the dissolved molecules into active (electrolytic) and inactive (non-electrolytic). Arrhenius gives nowhere in this memoir a precise account of the nature of the active and inactive portions. He indicates and illustrates what they might be, but that is all: he does not define. The most important special feature of the section is the statement that the active part increases on dilution. He introduces the notion of coefficient of activity, defined as follows: "The coefficient of activity of an electrolyte is the number expressing the ratio of the number of ions it actually contains to that which it would contain if the electrolyte were completely transformed into simple electrolytic molecules."

So far the considerations are purely theoretical; now comes an important step—their union with experimental data. Kohlrausch had shown that the molecular conductivity of an electrolyte was additively composed of two terms, one depending on the positive radical and the other on the negative radical. But in extremely dilute solutions of salts the value for negative radicals was nearly the same; therefore, according to Arrhenius, "the molecular conductivity of the active part of an acid (in dilute solution) is constant and independent of the nature of the acid," and as a corollary from this, "the better the (dilute) solution of an acid conducts electricity, the greater is its active part." For want of precise data for calculating the absolute value of the coefficient of activity, Arrhenius takes it as proportional to the molecular conductivity. Thus he is enabled to compare the activities of acids amongst themselves, and of bases amongst themselves. He finds at once that the activities of acids as thus determined from their conductivities agree well with our general notions regarding their strengths, and is led to the statement that "for acids and bases galvanic activity is accompanied by chemical activity."

Arrhenius proceeds to discuss double decomposition in electrolytic solutions, on somewhat hypothetical grounds, and arrives at a formula (containing coefficients of activity) which he applies practically to many important reactions. He gives, for example, a definite treatment of salt-hydrolysis based on the following principle: "What is common to all these cases is the necessity of regarding water as an acid (or as a base) which competes with other acids (or bases) present in the equilibrium." The theory is then applied to the displacement of one acid by another, to the influence of acid salts, and to equilibrium in heterogeneous systems. The consequence of the variation of the coefficient of activity in homogeneous and in heterogeneous systems is considered, and sections are devoted to the behaviour of molten electrolytes and to thermo-chemistry. He deduces the following important principle:

"The heat of neutralisation evolved by the transformation of a base and an acid, both perfectly active, into water and a simple salt is nothing but the heat of activation of water."

Nowhere does Arrhenius mention the word dissociation. The idea is, of course, there, but there is no identification of the "active part" of the electrolyte with free ions acting as separate molecules. The theory of electrolytic dissociation is, as it were, in solution; it has yet to be crystallised out.

This comprehensive memoir was presented to the University of Upsala in 1884 as a dissertation for the doctorate of the University. The impression made on the University authorities by Arrhenius's thesis was not favourable. In their view, apparently, it was neither very good physics nor very good chemistry, and only merited the award of a fourth class (*non sine laude approbatur*). After every allowance has been made for the novel and unusual character of the dissertation, it is difficult to see how the University of Upsala, the University of Bergman and Berzelius, should have condemned a brilliant thesis on the very subjects of affinity and electro-chemistry associated with these names. For the award amounted to a condemnation: in view of it Arrhenius could not normally become a docent in the University of Upsala.

Arrhenius sent copies of his paper to Clausius, Lothar Meyer, Ostwald and van't Hoff. "These celebrated men," he says, "with whom the Upsala professors were not to be compared, treated me as a colleague and not as a stupid schoolboy." Ostwald was eminently friendly, and handsomely acknowledged the priority of Arrhenius in work which intimately touched his own.

Oliver Lodge was impressed by the paper, and wrote for the Reports of the British Association in 1886 an abstract and critical analysis of it, extending to 30 closely printed pages. "The paper seems to me a distinct step towards a mathematical theory of chemistry. The title affixed to it is 'The Chemical Theory of Electrolytes,' but it is a bigger thing than this; it is really an attempt at *an electrolytic theory of chemistry.*"

In August, 1884, Ostwald visited Arrhenius in Upsala, and his visit had a marked effect on Arrhenius's future. Ostwald undertook to get Arrhenius appointed as a docent in Riga, and this offer no doubt led to the favourable reception of an application by Arrhenius for a similar post in physical chemistry at Upsala. The two men had, while together, projected a scheme of research on physical chemistry to be undertaken in Ostwald's laboratory in Riga, but the illness and subsequent death of Arrhenius's father kept him in Upsala. Through Edlund's influence he received in December, 1885, a travelling scholarship from the Academy of Sciences which enabled him to work in continental laboratories at discretion. The next five years were *Wanderjahre,* during which he worked with Ostwald, Kohlrausch, Boltzmann and van't Hoff. It was during this journeyman's period of his life that the theory of electrolytic dissociation was finally formulated.

His original paper had left the nature of the difference between the active and inactive portions of the electrolyte unsettled and the absolute value of the dissociation vague. As has been said, the theory was still in solution. The nucleus which determined its crystallisation came through van't Hoff's theoretical work on osmotic pressure and his interpretation of Raoult's experimental results.

Van't Hoff, in a memoir presented to the Swedish Academy on October 14, 1885, showed that it was possible to write for solutions an equation $PV = R'T$, analogous to the gas equation, where P, however, is the osmotic pressure instead of the gaseous pressure. The constant R′ was in many cases equal to the gas-constant, but in many others differed from it. Van't Hoff then wrote the general equation for dissolved substances in the form $PV = iRT$, where R is the gas-constant and i a coefficient sometimes equal to unity, but sometimes assuming values much greater, in particular for aqueous salt-solutions. For example, i is 1·98 for hydrochloric acid, 1·82 for sodium nitrate, and 1·78 for potassium chlorate. Van't Hoff contented himself with these empirical values and made no attempt at an explanation. His paper was published in 1886, but Arrhenius did not receive a copy until March, 1887. On the 30th of that month he wrote to van't Hoff from Würzburg: "Your paper has cleared up for me to a remarkable degree the constitution of solutions. If, for example, sodium chloride were normal in its behaviour—*i.e.*, if it consisted of simple molecules—its coefficient i would be equal to unity. But since i is much greater than unity, the natural explanation is to say that NaCl is partially dissociated, just as we say that at high temperatures I_2 is dissociated. Now this assumption might be deemed very rash were it not that on other grounds we are led to look upon electrolytes as partially dissociated, for we assume that they decompose into their ions. But as these ions are charged with very great quantities of electricity of opposite sign, conditions are such that we cannot in all cases treat a solution of NaCl as if it simply consisted of Na and Cl. The pressure on the walls cannot, however, be appreciably affected, so that in this case the solution acts as if Na and Cl were free. And when we consider which substances (according to Raoult's experiments) are abnormal, it is not the inorganic (*e.g.*, not $HgCl_2$, CO_2, H_2S, etc.), but the electrolytic (*i.e.*, substances which are conductors of the same order as salts), even when they are organic, *e.g.*, oxalic acid. . . . Since, according to the above assumption, electrolytes decompose into their ions, the coefficient i must lie between unity and the number of the ions. This in reality holds good, *e.g.*, for NaCl, KCl, KNO_3, NaOH, etc., which have two ions, the coefficient nearly reaches 2; for $Ba(OH)_2$, $CaCl_2$, K_2SO_4, etc., which have three ions, it almost approaches 3, and so on. . . . From the above assumption we can even calculate the value of i from the conductivity, and this I shall probably soon carry out; till now time has failed me. What I called in my paper 'Sur la conductibilité' active molecules are thus the same as dissociated molecules. One of the propositions which I then put forward would now be written: In all probability all electrolytes are completely dissociated at the most extreme dilution." Here we have the first appearance of the theory of electrolytic dissociation. If we cannot fix its birthday, at least we can its birth-month. It is clear, definite and concise, and all Arrhenius's previous theoretical treatment can easily be translated into terms of it. Van't Hoff accorded the new idea a favourable reception, but remarked that he had always thought of the dissociation into ions as being confined to an extremely small portion of the salt, though he confessed that he saw no grave difficulty in assuming a greater dissociation. Arrhenius replied: "It is true that Clausius only assumed that a minute quantity of a dissolved electrolyte is dissociated, and that all other physicists and chemists had followed him; but the only reason for this assumption, as far as I can understand, is a strong feeling of aversion to a dissociation at so low a temperature, without any actual facts against it being brought forward. In my paper on the conductivity of electrolytes, I was led to the conclusion that at the most extreme dilutions all salts would consist of simple conducting molecules. But the conducting molecules are, according to the hypothesis of Clausius and Williamson, dissociated: hence, at extreme dilutions all salt molecules are completely dissociated. The degree of dissociation can be simply found on this assumption by taking the ratio of the molecular conductivity of the solution in question to the molecular conductivity at the most extreme dilution." These two short excerpts give the gist of the complete theory.

Van't Hoff and Arrhenius now made their ideas available to a wider public by publishing them in the first volume of the *Zeitschrift für physikalische Chemie* in the latter half of 1887. Van't Hoff accepts Arrhenius's theory for electrolytes, and adds, finally, Avogadro's Law to those of Boyle and Charles as being applicable to dilute solutions. Arrhenius gives the relationship between van't Hoff's constant i and the degree of dissociation α in the form

$$i = 1 + (k - 1)\alpha,$$

where k is the number of ions into which the molecule of the electrolyte dissociates, *e.g.*, 2 for KCl, 3 for K_2SO_4. He compares the values of i calculated from Raoult's freezing-point data on the one hand, and from the molecular conductivity on the other, for some 80 different substances, and finds a very satisfactory accordance. In the second part of his paper he discusses the properties of electrolytes in aqueous solutions from the point of view of their additive character, which he attributes to the independence of their ions.

The theories of osmotic pressure and of electrolytic dissociation were now fairly launched, and, propelled by the driving-power of Ostwald through the waters of scientific opinion, they soon attained a world-wide recognition, though often meeting very heavy weather. That their reception was, on the whole, so favourable is indeed somewhat surprising, and must be attributed largely to their comparative simplicity. They could be easily tested practically, and little mathematics was required in their development, so that experimental work, centred originally in Ostwald's laboratory, but gradually spreading to others in Germany and abroad, was in the next decade assiduously carried out by a new generation of physical chemists. Their application by Nernst (1889) to electromotive force was an advance of the first order. Arrhenius himself played a principal part in the development. Amongst his important contributions to the subject published in Ostwald's *Zeitschrift* may be mentioned the theory of isohydric solutions (1888), the heat of dissociation of electrolytes and the influence of temperature on the degree of dissociation (1889), the conditions of equilibrium between electrolytes (1890), the hydrolysis of salts of weak acids and weak bases (1894), the alteration of the strength of weak bases by the addition of salts (1899).

During these years Arrhenius also worked on other physico-chemical subjects, for example, on viscosity of liquids and solutions, on conduction in hot gases and flames, on diffusion in aqueous solution, on the velocity of hydrolysis of ethyl acetate, and on the inversion of cane-sugar in acid solutions. In a paper on this last subject (1889), Arrhenius makes another theoretical contribution of great significance. He is discussing the effect of temperature on reaction velocity, which amounts at ordinary temperatures to an increase of 10 to 15 per cent. for one degree rise. This is much too great to be accounted for by increase of molecular velocity or diminution of viscosity, and the nature of the increase is altogether different from that exhibited in the temperature coefficient of ordinary physical properties. To explain the phenomena, Arrhenius assumes that the really active substance is not ordinary cane-sugar, but a hypothetica active variety in equilibrium with it. This " active " sugar is only present in very minute amount, but its proportion rapidly increases with rise of temperature. The influence of temperature is thus on an equilibrium rather than on a velocity, and van't Hoff's equilibrium equation relating temperature effect with heat of reaction (in this case heat of activation) may be applied. The resulting expression for the temperature coefficient has been confirmed for a large number of homogeneous and heterogeneous reactions, and although there is much that is arbitrary in its derivation, it is, in its general character, quite in accordance with modern ideas.

Returning to his personal fortunes, we find that after 1887 he was recognised abroad as one of the chief figures of physical chemistry, but the death of Edlund in 1888 deprived him of his stoutest champion at home, and greatly reduced his chances of obtaining suitable academic employment in Sweden. Arrhenius received a call to the chair of chemistry at Giessen in 1891, but, being intensely patriotic, he declined the offer on the chance of receiving an appointment as chief of the laboratory of physics in the Högskola (University College) at Stockholm, a post which at this time had become vacant. Arrhenius was successful in his candidature, and obtained this lectureship, which was in 1895 converted into a professorship, although once more against formidable opposition, only overcome by the strong backing of German physicists. During the years 1896–1902 he was Rector of the Högskola, and through his personality did much to stabilise and develop the struggling institution, notwithstanding that he had no fondness for administrative tasks. Although his laboratory was small and poorly equipped, the name of Arrhenius was sufficient to attract foreign workers, among whom may be mentioned Abegg, Bredig, Cohen and Euler, who afterwards succeeded him in the chair. Foreign distinctions also began to come his way. He was elected an honorary Fellow of the Chemical Society in 1898, and was awarded the Davy Medal in 1902. At last he received full recognition from his countrymen by the award of the Nobel Prize for Chemistry in 1903.

His interest had meanwhile been diverted from solutions to other fields of science, at first to cosmic and meteorological problems. With his friend the meteorologist, Nils Ekholm, he investigated the influence of the moon on the electric state of the atmosphere, on the aurora and on thunderstorms. In a long memoir (1896) he attempted to account for the onset and decay of glacial periods by the variation in the amount of carbon dioxide in the atmosphere. This gas exerts a selective absorption, allowing the solar radiation freely to pass inwards, but to a great extent stopping the lower temperature radiation from the earth outwards. Arrhenius calculated that from this greenhouse effect the temperature in the arctic regions might rise 8° C. if the carbon dioxide content of the atmosphere increased to somewhat more than double its present value, and that in order to get the temperature of an ice age between the 40th and 50th parallels, the value would have to sink to about half. The variation in the carbonic acid content he attributed chiefly to variation in volcanic activity. The problem of the Ice Ages is still a vexed question amongst geologists, but Arrhenius made a notable contribution to its discussion. Arrhenius published in 1900 another geological paper, namely, a theory of vulcanicity based on physico-chemical principles. According to it the sea-floor acts as a kind of semi-permeable membrane, permitting water molecules to pass but not silicate molecules. Water at no very great distance under the surface of the crust would be at a temperature above its critical point, and therefore a gas, and would be absorbed by the fluid magma under the great pressures existing. But by extrapolation from known data it may be shown that although at room temperature water is a much weaker acid than silicic acid, yet at high temperatures the reverse is the case, water at 1000° being probably 80 times stronger than silicic acid. In the magma, then, water will attack and decompose silicates, and thus be potentially stored up in the form of acid and base. When the magma on rising is cooled, the reverse process takes place; water is liberated, and at a certain height will overcome the pressure of the column above it, eject the superincumbent fluid, and cause a volcanic eruption. A volcano thus acts in much the same way as a geyser. The theory aims at explaining the proximate cause of eruptions, and has met with wide acceptance.

In 1903 he surprised his chemical friends by publishing his *Lehrbuch der kosmischen Physik*, a work of extraordinary learning and scholarship. In it he passes under review an extensive collection of observational material, and deals with it according to his own methods. The most striking novelty of treatment is the use he makes of radiation pressure, the existence of which had been predicted by Clerk Maxwell. It was applied by Arrhenius to various cosmic phenomena even before its experimental confirmation in the laboratory by Nichols and Hull, and by Lébedev. Arrhenius calculated that at the surface of the sun the repulsive force of the radiation would balance the sun's gravitational attraction on black particles of diameter about $1 \cdot 5\ \mu$ and specific gravity 1, and that smaller particles than these would be repelled. Schwarzschild made some necessary corrections, and showed that the maximum repulsion would be for completely reflecting particles (specific gravity 1) if their diameter was about $0 \cdot 16\ \mu$, and it would then be 10 times the gravitational attraction. From the sun, then, we might expect streams of such minute particles to be shot out in all directions. Many of these particles would be electrically charged from the ionisation existing in the sun's gaseous atmosphere. Arrhenius shows how the phenomena of the solar corona, comets the aurora, and the zodiacal light may be caused or influenced by these particles.

With the beginning of the present century Arrhenius's thoughts took a new turn. He became interested in serum therapy, and in the summer of 1902 he went to work in Ehrlich's laboratory in Frankfort. He collaborated with Thorvald Madsen and published jointly with him, in the same year, a memoir on the occasion of the opening of the Danish State Serum Institute, of which Madsen had been appointed director. It was entitled "Physical Chemistry applied to Toxins and Antitoxins." Madsen was responsible for the experimental methods, Arrhenius for the theoretical treatment. They maintained that the toxin-antitoxin combination (held by Ehrlich to be a firm chemical union) was in reality reversible, and governed by the ordinary mass-action law. The immunological phenomenon of antitoxin action was likened to the interaction of a weak acid and a weak base, such as boric acid and ammonia, which only partly neutralise each other. The work constitutes a classical study among the early researches into the underlying nature of immunity phenomena, and contributed to laying the foundation of "Immunochemistry," a term first applied by Arrhenius himself to a branch of biological research in which reactions of markedly specific character occur between biological principles of unknown chemical nature. Arrhenius pursued this type of research for a decade, and published two books dealing with it—"Immunochemistry" in 1908, and "Quantitative Laws in Biological Chemistry" in 1915.

In the year 1905 the Swedish Academy founded the Nobel Institute for Physical Chemistry, and of this new foundation Arrhenius was appointed director. It was housed at first in temporary quarters in Stockholm, but at Experimentalfältet, in a pretty park in the neighbourhood of the town, a small laboratory was erected, with an official residence attached. The laboratory was inaugurated in 1909. Here, with an assistant and a few research workers as guests, Arrhenius could work and write under ideal conditions on such problems of physical chemistry, physiological chemistry, immunochemistry, meteorology and cosmic physics as might please him.

The stormy period of Arrhenius's career was now definitely over, and from the time of his appointment to the Nobel Institute life went very smoothly with him. From being a scientific outcast in Sweden he became a scientific oracle, known and respected by all classes of the people.

He himself did little practical work in the new laboratory, but stimulated and encouraged others. One of his chief pleasures was to attend conferences in all countries for the purpose of meeting his scientific colleagues and discussing with them their special problems. He often visited England, and was elected a foreign member of the Royal Society in 1911. In 1914 he gave the Faraday Lecture to the Chemical Society, and the Tyndall Lectures to the Royal Institution. In the United States he delivered the Silliman Lectures n 1911.

Arrhenius liked to acquire knowledge at first hand, and visited many laboratories for this purpose. He spent, for example, three weeks in Rutherford's laboratory in Manchester, working at a practical course in radioactivity under Geiger. At the end of a week he had started a research on the solubility of active deposits, and was with difficulty dragged away from his electroscope to witness some of Jacques Loeb's starfish experiments at the Marine Biological Station.

He devoted a large part of the later years of his life to popularising science. A firm believer in progress through enlightenment, he sought to bring a knowledge of scientific fact and method before the general public. His clear and easy style made his books attractive, though the matters dealt with were often in themselves difficult. The first of these books, *Världarnas Utveckling* (Worlds in the Making), which treats in a popular manner some of the subjects of his *Kosmische Physik*, had an immediate and world-wide success, being translated into all the important European languages.

Happy in his work and happy in his domestic life, Arrhenius during his later years radiated contentment. He was twice married, in 1894 to Sofia

Rudbeck and in 1905 to Maria Johansson. By the first marriage he had one son, and by the second a son and two daughters.

His health remained good until the autumn of 1925, and although he made a remarkable recovery from the first attack of illness, he retired in February, 1927, from the Directorship of the Nobel Institute, when he was granted a full pension and the right to remain in the official residence. On October 2, 1927, he died, and was buried in Upsala on the eighth of that month after a solemn service in Stockholm on the previous day.

Arrhenius was of the old breed of natural philosophers, a true polyhistor, devoted to science at large. Being endowed with a memory both tenacious and accurate, he had a marvellous command of scientific fact. He was, however, no unimaginative empiric: his synthetic fancy played over the vast store of knowledge, and sought relations between apparently isolated regions. In consequence, his original ideas were concerned with borderland sciences—physical chemistry, geophysics, cosmic physics, immunochemistry. The conjunction in him of two faculties explains the character of much of his work—an aptitude for scientific speculation and an extraordinary facility in dealing with figures. He loved statistics, and it is recorded of him that as a very small boy he used to sit beside his father and help him in casting his laborious accounts. Arrhenius might begin a new line of work by the consideration of tables of numerical data by himself or others. He would frame a formula to fit them—an exercise at which he was uncannily expert—and then evolve a physical hypothesis to account for the formula. Or he might start with a bold speculation as to how two entities were related, formulate this relation, and check the formula by means of data of observation or experiment. There was constant interplay between the speculative and the quantitative sides of his mind.

Arrhenius had nothing academic about him save learning. In person he was stoutly built, blond, blue-eyed and rubicund, a true son of the Swedish countryside. His nature was frank, generous and expansive. He was full of robust vitality and primitive force. He had hearty likes and dislikes, and beneath his inborn geniality and good-humour was a latent combativeness, easily aroused in the cause of truth and freedom. He paid little regard to literature or art, but keenly appreciated natural beauty, especially the gladdening phenomena of spring. His lifelong interest in the lovely Northern Dancers had most likely an æsthetic as well as a scientific basis.

Sweden can boast of many eminent names in science, of which two are by common consent of the first magnitude—Linnæus and Berzelius. Since the death of Berzelius she has had no name to rank with these save that of Arrhenius.

J. W.

PAUL HEINRICH RITTER VON GROTH, 1843–1927.

By the death of Geheimrat PAUL HEINRICH RITTER VON GROTH, which occurred in Munich on December 2, 1927, mineralogical science has lost a leader who had been for a great part of his long life its most conspicuous figure.

His father, whose family resided in Kottbus, was a portrait painter who, after some residence in St. Petersburg and Rome, settled in Magdeburg, where Paul Groth was born on June 23, 1843. In the following year the family migrated to Dresden, where his childhood was passed. As a boy he showed an aptitude for drawing and water-colour painting, and was a studious reader, especially of books of travel. He was apparently inspired by the example of Alexander von Humboldt when, in 1862, he had to choose a career and decided to enter the Bergakademie at Freiberg; but he seems to have been attracted by mineralogy and crystallography even at the Gymnasium which he had attended from the age of 13 to 19.

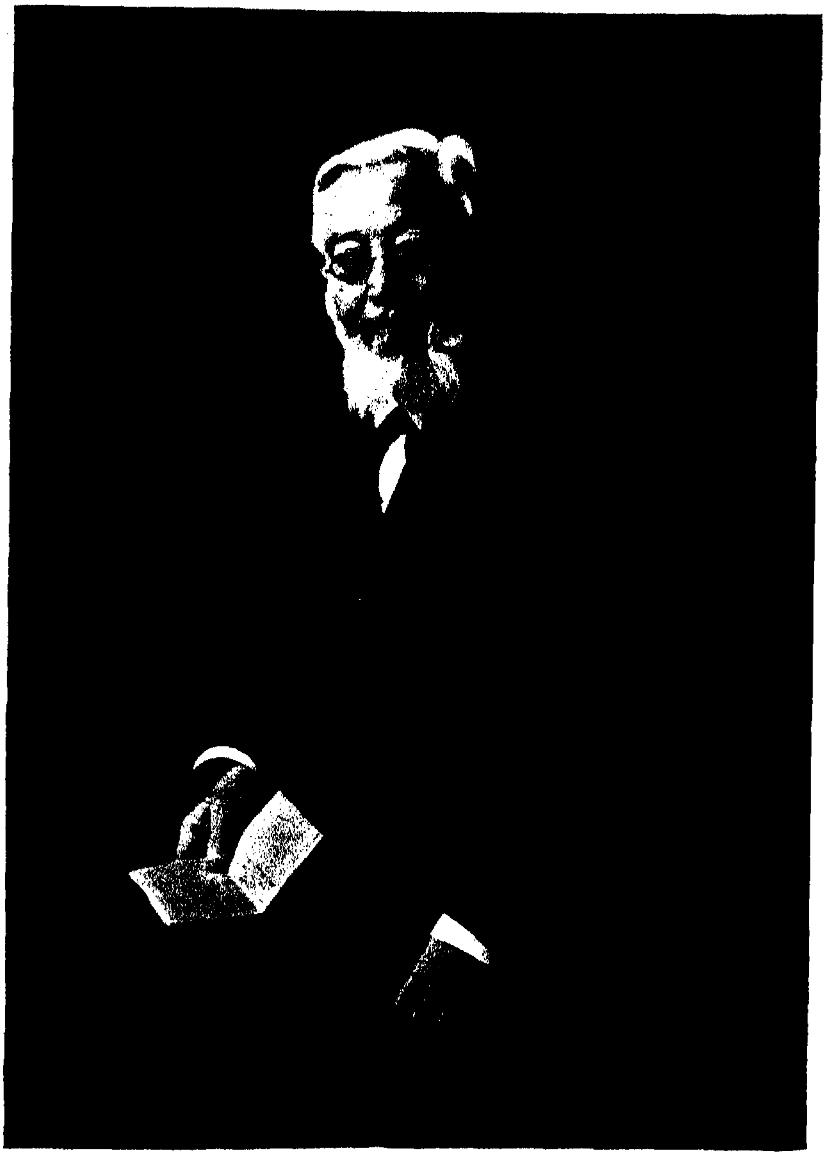
At Freiberg he followed the Mining Course and came under the influence of Breithaupt, the mineralogist, whom he always regarded as "the most inspiring teacher and most distinguished practical mineralogist since the time of Werner."

In 1864 he was studying analytical chemistry at the Polytechnische Schule in Dresden, and his earliest mineralogical publication (1866) related to a variety of titanite from the neighbourhood of Dresden to which Dana subsequently gave the name Grothite. In 1865 he migrated to the University of Berlin; there he interested himself in physics, particularly in relation to crystallography. His Inaugural Dissertation was on the isomorphism of the perchlorates and permanganates (1868), and in the same year he obtained his first teaching appointment as assistant to Magnus in the Physics Department.

From 1870 to 1872 Groth was Docent at the Bergakademie in Berlin, where his duties included general geology and led him to make geological excursions in the Eifel and in Rhineland; but his chief interest always remained centred in mineralogy and crystallography, especially in the physical properties of crystals and their relation to chemical composition.

He rapidly achieved such a reputation as teacher and investigator that in 1872 he was appointed to the Chair of Mineralogy in the newly constituted University of Strasbourg, where he remained for eleven years. Here he was provided with adequate funds, and soon created a finely equipped institute with a notable collection of minerals admirably arranged for teaching purposes; it was not long before students from many countries began to seek his laboratory. Overwork during the first year led to temporary lung trouble, the only illness of his long life.

From this time onwards his career was one of unceasing activity as a teacher



Prof. P. Groth

and director of research, and was marked by a succession of important publications. In 1874 he published his 'Tabellarische Übersicht der einfachen Mineralien'—a comprehensive list of the mineral kingdom, containing not only a systematic classification of species, but also a critical survey of views on their chemical composition; subsequent editions with much new material appeared in 1882, 1889, 1898, and were followed in 1921 by a new survey, 'Mineralogische Tabellen,' in conjunction with Mieleitner. In 1876 he published his famous 'Physikalische Krystallographie,' a most readable and suggestive treatise which was for many generations of teachers and students an attractive introduction to a science that had previously been presented in a very unattractive form. Sir Lazarus Fletcher has recorded the fact that he was led to take up the study of the subject by happening to see a copy of the book in that year. Subsequent enlarged and revised editions appeared in 1885, 1895, 1905.

In 1877, Groth started the first volume of the *Zeitschrift für Krystallographie und Mineralogie*, which became universally known as *Groth's Zeitschrift.* This he edited with great skill for thirty-nine years, enlisting the co-operation of a large number of mineralogists from all countries; it was conspicuous for its international character and for the value not only of the original papers, but also of the abstracts which it contained. In a memoir which appeared shortly before his death in the *Zeitschrift*, he gave an account of its inception and a history of its progress during the period of his editorship. The fifty-eighth volume (1923) was a *Festband* in honour of his eightieth birthday and contained papers by many of his old pupils, representing twelve countries.

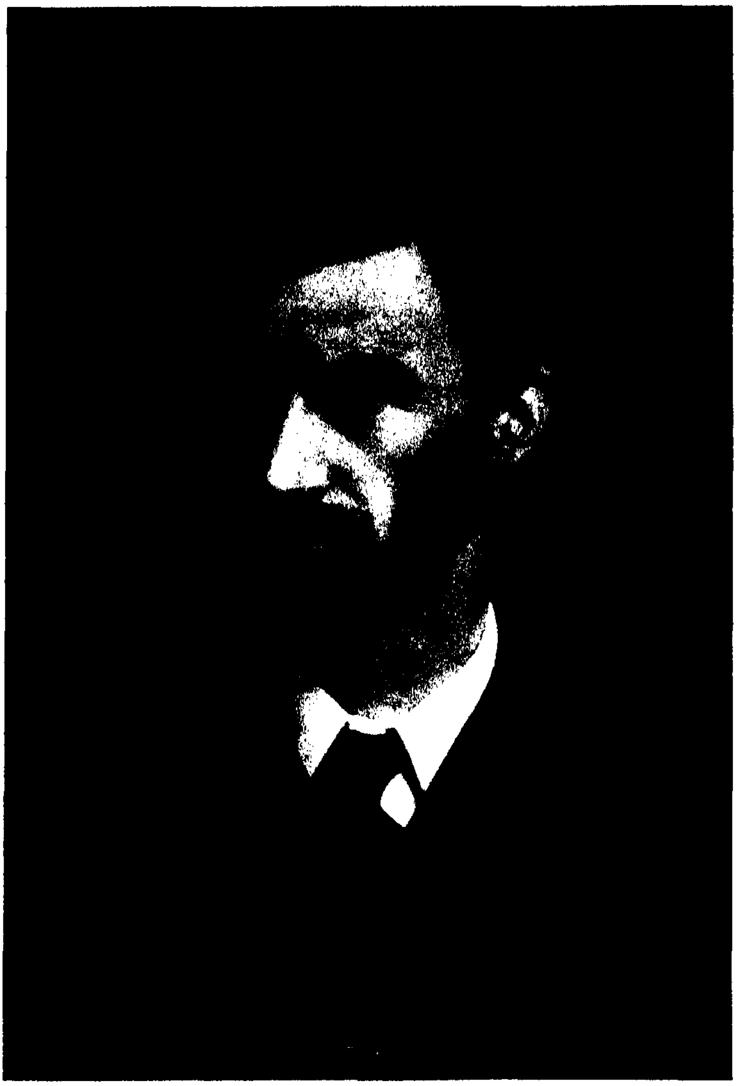
In 1878, Groth's 'Mineraliensammlung der Kaiser-Wilhelms-Universität Strassburg' not only showed how active he had been in getting together an important teaching collection in six years, but also served as a model for future mineral catalogues. His reputation as the foremost teacher of mineralogy and crystallography had been firmly established when, in 1883, on the death of Kobell, he was transferred to Munich as professor of mineralogy and Director of the State Collections. His department at Strasbourg had been remarkable for its international character; young men from all the countries of the world, who were destined to be the professors of the next generation, flocked to his laboratory; and at Munich the new department preserved the same character. Here again he inspired a remarkable succession of researchers, organised a fine institute of mineralogy and crystallography, and by uniting several old collections built up a worthy State Museum especially rich in Alpine minerals. A guide to the collection appeared in 1891, an introduction to the study of precious stones in 1887, an introduction to chemical crystallography in 1904, the elements of physical and chemical crystallography in 1921. But his energies were mainly concentrated on a sequel to his life's work in the form of a vast book of reference—a complete dictionary of the physical and chemical characters of crystallisable substances—mainly based upon the material which he had collected in fifty-five volumes of the *Zeitschrift*. 'Chemische Krystallographie' appeared in five volumes between 1906 and 1919, and is an invaluable work of reference, giving full data for more than 7,000 substances.

A life so strenuously devoted to teaching, to the writing of treatises and text-books, to the organisation of university departments, to the inspiration of students, and to the supervision of their investigations, left little time for research. Except through the work of his pupils, the name of Groth is not associated with any great discovery. But he was always occupied in welcoming and in fostering new ideas, and he always had problems for his students to work out. At the beginning of his career he made a striking study of changes of form in the crystal produced by the substitution of one element or radical for another, especially among organic compounds, and introduced the term "morphotropy" for this conception; it guided his masterly studies on the composition of minerals; the connection between chemical composition and crystalline form was always his chief interest. The avidity with which, in his old age, he absorbed the new X-ray experimental work and modern theories about crystal structure is evidenced by his latest publications. The last of these, 'Entwickelungsgeschichte der mineralogischen Wissenschaften' (1926), and 'Vorgeschichte, Gründung und Entwicklung der Zeitschrift für Kristallographie in den ersten fünfzig Jahren' (1927) were written after his retirement from active teaching and when his eyesight had almost failed.

Devotion to academic duties and to the *Zeitschrift* left comparatively little time for travel, but Groth made mineralogical expeditions in France and Italy. In 1893 he visited Canada and the United States, and in 1897 (Geological Congress in St. Petersburg) travelled in the Urals. In 1904 he visited England for the British Association meeting and made a tour in Wales. In Great Britain his position in the scientific world was fully appreciated fifty years ago, when he was elected an honorary member of the Mineralogical Society. He was afterwards elected foreign correspondent (1895) and foreign member (1900) of the Geological Society (and received its Wollaston Medal in 1908), foreign member of the Royal Society in 1911, and honorary fellow of the Mineralogical Society of America (1926). He received the Hon. Sc.D. degree from Cambridge in 1904, and was also an honorary graduate of Geneva and Prague.

The present writer has a vivid recollection of Groth as an active young professor at Strasbourg in 1881; energetic and busy with his class and his collection, but always accessible and ready to help and inspire the little group of eager research students who worked in his laboratory. He always impressed them as a strenuous worker, a great teacher, and a most genial, sympathetic, and loyal friend.

H. A. M.



W H Dines

WILLIAM HENRY DINES, 1855-1927.

WILLIAM HENRY DINES was elected a Fellow of the Society in 1905, when he was in his fiftieth year. He had been President of the Royal Meteorological Society in 1901 and 1902. Devoted to the dynamical and physical aspects of meteorology, he had made himself a master of anemometry, having designed and brought into use a number of instruments. He had personally taken the lead in the exploration of the upper air of the British Isles by means of kites. The results of soundings off the west coast of Scotland from a tug chartered for the purpose had been described in the Philosophical Transactions of 1903.

From that time onward the history of his life is that of his contribution to dynamical and physical meteorology. Kites which made use of miles of invisible steel wire became " taboo " inland when motors took possession of the roads. Teisserenc de Bort, who employed kites on an elaborate scale in the neighbourhood of Paris, had the experience of holding up a locomotive on the railway, a steamer on the river with a tangle of steel wire, and of upsetting the unsuspecting occupant of a skiff. Dines was careful enough to avoid any serious incidents during the ten years of his using kites. His train did not run to the six or seven on a single trunk wire which may be necessary to reach the greater heights ; but even in the most carefully selected spots in this country, kite-flying without any flotation became too dangerous by 1912. Moreover, by that time free balloons carrying self-recording instruments had come into use, and the identification of the stratosphere beyond the reach of thermal convexion offered a new attraction for scientific endeavour. Dines continued to devote himself to the exploration of the upper air at his house at Pyrton Hill, Watlington, until the end of 1913, and thereafter until 1922 at Benson, Oxfordshire, in association with the Meteorological Office. He supplemented the work by the study of solar and terrestrial radiation, amongst other things, until his death, which occurred on Christmas eve of 1927.

For taking up meteorology as a scientific pursuit there was abundant precedent ; but Dines joined at a time when the subject was beginning to feel the effect of the cleavage between physics and meteorology (or geophysics generally), which resulted from the development of the weather-map on the one side, and of electrodynamics on the other, in the middle period of last century.

He was born on August 5, 1855. His father was George Dines, who died in 1888, the inventor of a hygrometer, well known to students of physics, and an active Fellow of the Meteorological Society. Through his mother, Louisa Sara Coke, he could claim descent from the Cokes of Norfolk. Of his family, he was the only son who survived childhood : two older sisters, who lived to grow up, pre-deceased him. Never at any time exactly robust, as a child he was delicate :

he was educated at Woodcote School, Windlesham. He learned engineering as a pupil at the Nine Elm Works of the South Western Railway, and after completing his apprenticeship there went to Cambridge, with little knowledge of mathematics, or, indeed, of any of the more academic subjects. He joined Corpus Christi College, which at that time had not yielded to the fashion for awarding scholarships before residence. He took the B.A. degree in 1881 as twentieth wrangler. He remained at Cambridge as a mathematical "coach" for a year. In 1882 he married Catharine Emma Tugwell, the daughter of a London clergyman, whose brother subsequently became Bishop of Equatorial Africa. He continued to teach mathematics, first as assistant to an army coach, and subsequently in correspondence classes; but as years went on he allowed his meteorological work to absorb his whole time.

He carried away from Cambridge an unfavourable impression of College officialdom: he held resolutely to the opinion that a self-respecting graduate could not apply for the higher degree of M.A. so long as it could be acquired simply by payment. He was habitually modest, and must have failed to realise the self-importance that is natural to a graduate who can produce twenty-five pounds in cash within three years of his first degree. Notwithstanding his austere opinion, he kept up his connection with Cambridge by sending both his sons there in their turn: he was himself responsible for their preliminary education.

About the time of taking his degree the Meteorological Council, which in 1876 had replaced the Meteorological Committee of the Royal Society of 1867, had embarked upon a new policy of supplementing the work of the Meteorological Office by "special researches" upon the atmosphere, in which W. J. Russell for fog investigation, Capt. W. de W. Abney and Sir G. Stokes for clouds, Capt. Andrew Noble for gunnery, Capt. Templer for ballooning, and others took part. This was in amplification of the plan, due primarily to Galton, of using the continuous records of various physical elements at the observatories to explain the phenomena of weather. The effort languished in course of ten years without much permanent achievement, except the Stokes' sunshine recorder. One of the proposed researches was the measurement of humidity, which was of considerable interest to the elder Dines, and among the first objects of W. H. Dines's attention was an apparatus for determining vapour-pressure in absolute measure. It brought the present writer into contact for the first time with both father and son, and led to a life-long and intimate friendship.

The discussion which followed the Tay Bridge disaster of December 28, 1879, when a train, crossing the bridge in the darkness of the winter evening, was carried bodily away with the bridge into the river by a squall of wind, claimed the attention of meteorologists as well as engineers in order to decide what allowance ought to be made for wind-force on engineers' structures.

In consequence, Dines became the most active member of a Wind-Force Committee appointed by the Meteorological Society in November, 1886, "to investigate the relation between Beaufort's Notation of Wind-Force and the equivalent velocity in miles per hour, as well as the corresponding pressure in pounds per square foot: to inquire whether any existing scale can be adopted or modified, and, if not, to determine such equivalents as can be recommended for general and international use: also that the Committee be requested to report on the best mode available for the attainment of a satisfactory solution of the entire question of wind-force."

Considering the amount of attention that was given to anemometry in the first two-thirds of the nineteenth century, it is remarkable that "the question of wind-force" was so imperfectly represented at that time. A table "To find the force of the wind" in Molesworth's Pocket-book, dated 1891, gives $P = 0{\cdot}00492\,v^2$ as the relation between force P in pounds per square foot and the velocity v in miles per hour. Dines changed the factor to 0·003 for the force upon a disc one square foot in area. The subsequent measurements at the National Physical Laboratory in connection with aëronautics have only slightly modified the simplicity of the relation. The transition from three "significant figures" with a tolerance of 40 per cent. error to a single figure with very little error at all is quite typical of Dines's attitude towards the dynamical and physical problems of meteorology. Scrupulously accurate himself, he always kept the practical accuracy of measurements in mind.

In the course of his investigation he reduced the factor for the ratio of the run of the wind to the run of the cups of the Robinson anemometer from 3, which Dr. Robinson had adopted on what are called "theoretical grounds," to 2·1 for the standard pattern of nine-inch cups with two-foot arms, and recommended the special determination of the factor for instruments that were not of the standard pattern.

For these determinations he used a whirling arm driven by a steam engine in the garden of his father's house at Hersham, in Surrey, taking all precautions to avoid the errors that so easily beset a measurement of that kind. With the same apparatus measurements were obtained of the "constants" of a number of different kinds of anemometer, including the tube-anemometer, and the bridled anemometer, a modification of the Robinson anemometer designed by Sir G. G. Stokes, which rivalled the tube-anemometer in the effective representation of gusts. The "tube-anemometer" was a new departure of Dines's own design, which gave a record of velocity corresponding with the transient variations of the pressure of the wind on the opening of a tube. It is now commonly known as the pressure-tube anemograph: the original is in the Science Museum, where also may be found Dines's apparatus for illustrating the formation of a tornado cloud, and his instruments for the exploration of the upper air. The manufacture of the anemograph was entrusted to R. W. Munro, now of South Tottenham. It is of the type which is known in modern engineering practice as the Pitot tube, an instrument devised in 1732 to measure the flow of water. It required, however, special arrangements for exposure to the wind, which varies in direction as well as force, and a special contrivance of a shaped float to transform the pressure into a velocity scale.

As a meteorological instrument, it was completed by an additional contrivance for recording the direction as well as the velocity which, however, was not added to the instrument as placed on the market until the Meteorological Office asked for it. As then completed, the instrument provided a record of wind which is indispensable for the study of dynamical meteorology.

On the Wind-Force Committee Dines was associated with Mr. Richard Curtis, who was in charge at the Meteorological Office of the automatic records received from the observatories; and in continuation of the investigation of the problem of wind-force a number of instruments were set up for the Meteorological Council on Salt Island, Holyhead, where a free exposure was available: the instruments included the standard Robinson, Stokes's bridled anemometer, the pressure-tube and a pressure-plate. A Robinson instrument and pressure-tube were also established on St. Mary's, Scilly, and subsequently elsewhere. In these operations, informally, Dines was honorary consulting engineer.

The pressure-plate included in the equipment at Holyhead was of his own design, to avoid the difficulty in the use of instruments of that kind which arises from the inertia of the moving parts. A chain passed over a pulley behind the plate, and was kept taut by a sliding weight in the supporting column, which carried a ratchet. If a gust pressed the plate the ratchet held it, and any gust that followed could only move it further if it exceeded the original gust in force. The ratchet was released each day by the attendant, and so the maximum reached each day was indicated. The behaviour of the instruments formed the subject of a note in the annual reports of the Council, but in the end they merely demonstrated the efficiency of the pressure-tube.

The expression of wind in terms of velocity instead of force is perhaps a meteorological curiosity. For engineers the force may be regarded as a more important consideration than the velocity, and the construction of the pressure-tube to give velocity instead of the force, which is, in fact, the operative element, involves elaborate design. But the whole structure of dynamical meteorology is raised upon velocity rather than force, and it is doubtful if any attempt has been made to arrange it otherwise.

For many years the pressure-tube-anemometer was regarded as a curiosity rather than a working instrument in meteorological observatories, because the practice of publishing hourly velocities had become established in international practice, and the pressure-tube gave a continuous record containing gusts and lulls of which international meteorology took no account, and which could neither be integrated nor averaged. The record at Richmond, for example, of a squall that uprooted part of the chestnut avenue at Bushy Park, is included among a selection of records in the Report No. 9 of the Advisory Committee for Aeronautics, very informing as it stands; it loses its significance if integrated over an hour. It is, in fact, not exactly possible to give an adequate numerical description of the record.

In the investigation of the upper air Dines was equally successful. His memoir on the structure of the upper air, drawn up for the Meteorological Office with the title, "The Characteristics of the Free Atmosphere," and published in 1919, is unrivalled as a concise and effective summary of many years of enterprise. In this, as in the investigation of wind, a Committee was necessary, but Dines was the prime mover. A committee may act as a lubricant, or it may keep a sand-box and work the brake, but for a scientific enterprise most committees are ineffective without a prime mover. In this case the Committee was a joint one appointed in 1901, in part by the British Association with a grant of £75, and in part by the Royal Meteorological Society with a grant of £25. The funds were supplemented by £75 from the Government Grant Committee.

Dines suggested the inquiry first to the Meteorological Office in 1900; but at the time the Meteorological Council had nothing to offer, except instruments for a base station; it was in extreme poverty. It was called upon by superior powers to provide the funds necessary to pension its chief officer and its senior staff by economies in its expenditure: its work was in arrear, and members of its staff were adopting various expedients to make ends meet. Its application to the Royal Society to obtain a special grant of £500 for work on the upper air was merged in a Treasury inquiry which originated in a miscalculation. The Scottish meteorologists had been induced to expect that the transformation of Kew Observatory into the National Physical Laboratory would set free funds that might be diverted to the support of the observatory on Ben Nevis. It was not until 1905, when the agitation prompted by the miscalculation had resulted in a reconstitution of the office, that the sum of £500 a year, in lieu of the other contributions for the support of the investigation of the upper air, could be spared from the public funds allocated to the Meteorological Office.

The results have more than justified the expenditure. They are set out in various reports of the Meteorological Office, with papers in the 'Philosophical Transactions,' the 'Quarterly Journal of the Royal Meteorological Society,' and elsewhere. The 'Journal of the Aeronautical Society of Great Britain' ought to be mentioned also, because the association with that Society, then mainly in charge of Major Baden Powell and Mr. Eric Stuart Bruce, was very close. Dines was a member of the council, and after the war was elected an Honorary Fellow of the reorganised Society.

The investigation of the upper air, which was one of the first items in the project of the Meteorological Office in 1877, was brought into the programme of international co-operation in 1896 by a fourfold appeal: first the development of scientific ballooning in Germany, the results of which are recorded in three splendid volumes of the Aeronautical Society of Berlin, edited by Richard Assmann; secondly, the establishment of the Blue Hill Observatory in the United States by Abbott Lawrence Rotch; thirdly, the foundation of the observatory for kites by Teisserenc de Bort at Trappes, and fourthly, the adoption of the method of exploration by sounding balloons carrying recording instruments by Assmann, Hergesell and Teisserenc de Bort, initiated by Hermite and Besançon. Hugo Hergesell, of Strassburg, became president of an International Commission appointed in 1896.

Dines began in 1902 with kites at Oxshott, and supplemented the results by similar observations during the summer in the West of Scotland from a small island in Crinan Bay, with the use of the tug *Countess*, chartered for the purpose in July and August. In the next summer another steamer, the *Renown*, was chartered, and was found much less satisfactory: in 1904, through the good offices of Admiral Sir W. Wharton, Hydrographer of the Navy, H.M.S. *Seahorse* was lent by the Admiralty.

The kite results at Oxshott were embodied in an official memoir by E. Gold on "Barometric Gradient and Wind-Force in 1908." One of the striking results of the observations at Crinan was the shape of a cloud-canopy and the evidence of the difference in the height of cloud on a hillside and overhead, and the consequent liability to error of personal estimates of cloud-heights.

From the beginning Dines used diamond-shaped box-kites of his own design, at first with meteorographs by Richard Frères, of Paris, supplemented by a device of his own, consisting of two glass tubes, dipping one into water and the other into alcohol in order to determine the maximum height of an ascent without risking an expensive instrument. "I hope to perfect this method, since there are many occasions on which a kite and a couple of glass tubes might be risked when one would hesitate to send up instruments costing £20. The apparatus above described is now in use every day when the wind is suitable . . . a velocity of 15 miles per hour is necessary."

Subsequently a special meteorograph for use with kites was designed to record on a flat board pressure, temperature, humidity and wind. The board with the recording pens was covered by a linen case. After it had been in use for some fifteen years, service officers who wanted a recording instrument for kite-balloons were quite put out by its "home-made" appearance, and reported it as the trial model of an inexpert designer.

In 1906 Dines thought Oxshott had become too populous, and found a secluded spot on the northern slope of the Chiltern Hills, near Watlington, where he could carry on his researches with the assistance of the Office. From the modern point of view it may be regarded as an unusual arrangement. Modern administration tends to insist that anyone in charge of research shall make a precise statement of what he is going to discover before he is allowed to look for it, a limitation which W. H. Dines would have found irksome. He rented the house, and moved there in November of that year, and began to develop the use of sounding-balloons, maintaining the investigation with kites at the same time. A feature of the new arrangement was the employment of a mechanical assistant, H. W. Baker, who became a permanent feature of the establishment. A professional assistant was added in 1914, and a second mechanical assistant at a subsequent date. A letter to *The Times*, in 1903, pointing out that kite-flying at sea would add to the pleasure and excitement of a yachting cruise, brought in Mr. C. J. P. Cave, of Ditcham Park, who set up kite-gear on his estate. He co-operated also in the ascents with sounding-balloons, and introduced into this country the practice of using small pilot-balloons to determine the currents in the upper air which had been developed on the Continent especially by A. de Quervain, and has now become an almost universal practice to obtain information for aviation. "Dr. G. C. Simpson, who had recently concluded a year of work as voluntary assistant at the Office, and had taken up the duty of lecturer in meteorology in the University of Manchester, made preparations for an experimental station at Glossop Moor, in Derbyshire, having previously spent three weeks kite-flying from a trawler on the North Sea. Out of this has been gradually developed the present complete equipment for the investigation of the upper air during the years 1908–9 at the Howard Estate meteorological station at Glossop Moor, with suitable provision for experiments with kites, captive balloons, pilot-balloons and registering balloons in connection with the Physical Laboratory of the University of Manchester." Other persons co-operating were Mr. S. H. R. Salmon, with a kite-station on the downs at Brighton, Captain C. H. Ley, R.E., with sounding balloons at Ross in Herefordshire, and subsequently at Bird Hill, Limerick. In later years Mungret College, Limerick, became a regular station for work in the upper air.

In this way organised investigation was undertaken successfully on the days appointed by the International Commission.

For the sounding balloons a special instrument was designed. It carried no clock, but recorded temperature in relation to pressure. Duplicates for use at the other stations were made and calibrated in Dines's workshop. The cost worked out at £1 each, in comparison with £15 to £20 then required for instruments used in other countries, a great advantage in consideration of the risk of loss of instruments carried beyond the coast. Further economies were secured by the lightness of the instrument which, with the case, did not weigh more than 2 ounces (60 grammes): in favourable circumstances a single, comparatively small balloon can raise the instrument to 20,000 metres.

The first successful ascent with apparatus designed and made at Pyrton Hill was obtained by Mr. Cave on July 1, 1907, and between that date and December 31, 1908, 133 balloons had been sent up in the British Isles, and of these 82 had been recovered." It was a disappointing proportion compared with continental countries.

One striking result of the moderate price of the equipment was the launching of balloons at hourly intervals from Manchester under Prof. Petavel's direction, and with Prof. Schuster's support on two occasions, June 2 to 3, 1909, and March 18 to 19, 1910. Dines's meteorograph is used at present not only in this country, but also in Canada and in Spain. A modification of it is used also in India.

As the use of sounding balloons and pilot-balloons developed the use of kites was restricted. In 1910 J. S. Dines, younger son of W. H. Dines, was appointed to conduct an investigation of the upper air with pilot-balloons, and the special study of winds for the Advisory Committee for Aeronautics; pending the completion of buildings at the new Aircraft Factory at Farnborough, he was accommodated at Pyrton Hill. The connection brought Mr. Dines into close association with the meteorological side of the work of the Advisory Committee.

At the end of 1913 Pyrton Hill was reclaimed by its owner, and Dines was compelled to seek accommodation elsewhere. By that time the scientific value of the use of sounding-balloons had become emphasised. Dines accordingly sought the best locality for the purpose. With that object in view, and with the assent of the Meteorological Office, he acquired the property at Benson, to which he moved early in 1914. We had planned various schemes of investigation in connection with the Office, but the war intervened. During the war the co-operation of the various centres of meteorological activity at home was consolidated by the Meteorological Sub-Committee of the Advisory Committee for Aeronautics, which afforded periodical opportunity for personal consultation. Benson was active both in experimental work and in the co-ordination and discussion of results. After the war Mr. L. F. Richardson joined Mr. Dines and Benson became a centre of notable scientific activity; but the full programme of co-operative work on the upper air was not renewed. Mr. Dines resigned the official charge of the observatory as from June 30, 1922. His elder son, L. H. G. Dines, was transferred from Valencia to Benson as Assistant Superintendent to take charge of the official work, which continued at Benson until September 30, 1923, when it was transferred to Kew Observatory. The provision is almost humble, compared with what is thought "worth while" in Germany, Italy, Holland or the United States.

One of the subjects of the joint activity of Dines and Richardson was solar radiation. Dines seems to have pondered over it in the early years of the war He endeavoured to make out the particulars of a balance sheet of the radiation which is received as short waves from the sun, and, after certain transformations in the atmosphere or at the surface, is subsequently dispersed in space as long waves. He presented his draft balance-sheet to the Royal Meteorological Society in November, 1916, and, relying on the Smithsonian Institution for measures of the incident solar radiation, turned his attention to the radiation from the atmosphere (called "the sky" in literature of the subject) in relation to the radiation from the earth. The subject is very intricate. Dines writes: "Abbot and Fowle give the effective radiative temperature of the earth as about 255°, but this value is obtained by ignoring 37 per cent. of reflected sun-heat. Since the atmosphere reflects more than one-third of the radiant heat that falls upon it, the earth cannot be called a black body, and the term 'effective radiative temperature' applied to it is apt to mislead the unwary."

By 1920 he had designed an ether differential radiometer by which to compare the radiative temperature of a tank of water with that of a portion of the sky. Ether in contact with its vapour was used as a sensitive thermometer. The instrument was used up to May, 1920, and then a battery of thermojunctions, copper-eureka, was used instead. The junctions were mounted near the end of a horizontal tube, which was inserted through the side of a water tank, and radiative communication with the sky was established by a silver mirror, which could be adjusted for communication with any part of the sky or ground.

The apparatus works with the facility of a searchlight, and brings within the possibility of solution a large number of problems in the balance-sheet of solar and terrestrial radiation. Readings of the relation of radiation from earth and sky about sunset have been summarised in the 'Meteorological Magazine' quarterly since the beginning of 1921, and Dines's last paper, which his elder son shared, put together the results obtained. A model of the instrument was exhibited at the British Empire Exhibition, but unfortunately no provision has yet been made for its manufacture.

During 1927 Dines's health gradually failed. News of his death on Christmas eve only reached London on the day of his funeral in the churchyard of Benson on December 28.

The three subjects in which his work represents determined and coherent effort are mentioned particularly, namely, wind-measurement, investigation of the upper air, and solar and terrestrial radiation; but his activity was by no means confined to those subjects: he was equally at home in the design and construction of instruments, the co-ordination and discussion of results, and the consideration of current generalisations and theory. He transformed the wheel-barometer into an instrument of precision that may be seen any day in the entrance-lobby of the Meteorological Office at South Kensington, or at any one of the chief observatories of the Meteorological Office. At my own request he produced a most ingenious instrument, which recorded only rapid variations of the barometric pressure. He had also a sunshine-recorder which few people have seen, and a temperature-recorder which is in use at Southport Observatory.

In his co-ordination of the upper air results he followed the most approved statistical methods, and even ventured to differ from some of the most prominent authorities on that embarrassing subject. He would deal with a question of weather in relation to health quite effectively.

He seemed to take a kind of austere pleasure in noting the generalisations of less perspicuous meteorologists that would not bear the test of rigorous examination: he would not allow that cold winter weather was a necessary characteristic of anticyclonic weather, nor that anticyclones always meant fine weather—indeed, "anticyclonic gloom" is always associated, by those who know, with Dines's name.

He was of very retiring and unassuming disposition, and would not even allow other people to put him forward. He was a member of the International Commission for the Exploration of the Upper Air, but not a regular attendant at the meetings. He was probably as uncommercial a philosopher as ever existed. His first step after completing an instrument was to have it publicly exhibited, and so rescued from the wiles of the patentee, a practice that instrument makers neither understand nor appreciate. If one asked him about the design of an instrument, he might produce one from his workshop, saying that he had made it, but people did not seem much interested in it. His latest radiation instrument may prove to be the most powerful weapon that physical meteorology has in its armoury, but it is not on the market. The only "honours" that found their way through his seclusion besides the Fellowship of the Royal Society was the "Symons" Medal in 1914, and the Buchan Prize in 1924 from the Royal Meteorological Society.

He was never in a hurry, but never unemployed. "Ohne Hast und ohne Rast" is a true description of his career. A very loyal friend, an invaluable ally, and, through all, a companion with the inexpressible charm of complete sincerity.

His two sons are in the Meteorological Service of the Air Ministry. The younger joined the office staff in 1907 as student assistant after a year's experience of his father's work at Pyrton Hill. He was the first official in charge of meteorology at South Farnborough, and is now Superintendent of the Forecast Division. The elder, after serving as an engineering pupil at Swindon, was at Eskdalemuir, subsequently at Valencia, and is now Assistant Superintendent in charge of the Upper Air Section at Kew Observatory, and in that sense is the successor of his father, though the conditions are a little different.

NAPIER SHAW.

INDEX TO VOL. CXIX. (A)

Adam (N. K.) The Structure of Thin Films, XI, 628.
Alkali metal vapours and the Saha theory (Wilson), 1.
Alkali metals and their alloys, spectra (Walter and Barrett), 257.
Arrhenius (S. A.) Obituary Notice, ix.
Aston (G. H.) *See* Ellis and Aston.
Atkinson (R. d'E.) Statistical Experiments on the Motion of Electrons in Gases, 335.

Barrett (S.) *See* Walter and Barrett.
Beta-rays, photographic action and velocity (Ellis and Aston), 645.
Bickley (W. G.) The Influence of Vortices upon the Resistance experienced by Solids moving through a Liquid, 146.
Bismuth crystals, specific resistance in strong magnetic fields (Kapitza), 358, 387, 401.
Boys (C. V.) Solid Diploidoscope Prisms, 489.

Cathode rays, diffraction, II (Thomson), 651.
Cathode rays, diffraction by thin celluloid films (Reid), 662.
Cathode rays, diffraction by thin metal films (Ironside), 668.
Cathode ray oscillograph, overvoltage measurements (Newbery), 680.
Cathode rays, scattering (Schonland), 673.
Catalysts, film, method of measuring surface area (Constable), 196, 202, 710.
Chalkin (F. C.) *See* Richardson and Chalklin.
Chapman (S.) On the Brownian Displacements and Thermal Diffusion of Grains suspended in a Non-Uniform Fluid, 34. The Molecular Displacements in Diffusing Gas-Mixtures, 55.
Chapman (S.) *See* Price and Chapman.
Chemical change, mechanism, I (Lowry and Owen), 505, 710.
Cockcroft (J. D.) On Phenomena occurring in the Condensation of Molecular Streams on Surfaces, 293.
Compton effect photometer (Raman and Sogani), 526.
Condensation phenomena at different temperatures (Powell), 553.
Conductivity, electron theory and the kinetic method (Nordheim), 689.
Constable (F. H.) A New Interference Method of Measuring the Surface Area of Film Catalysts. I—Theory, 196, 710. II—Nickel, 202.
Cook (S. S.) Erosion by Water-hammer, 481.
Copper, spark spectrum (Menzies), 249.
Cubic surfaces, nodal (Timms), 213.
Cyanite, structure (Taylor and Jackson), 132.

Davies (L. P.) The Photo-Electric Properties of some Metals in the Soft X-Ray Region, 543.
Davies (S. J.) and White (C. M.) An Experimental Study of the Flow of Water in Pipes of Rectangular Section, 92.

Diffusing gas-mixtures, molecular displacements (Chapman), 55.
Diffusion of grains in fluid, brownian displacements (Chapman), 34.
Dines (W. H.) Obituary Notice, xxiii.
Dispersion, anomalous rotary, of sodium tartrate (Lowry and Vernon), 505.

Edwards (R. S.) The Effect of Temperature on the Viscosity of Neon, 578.
Electron emission in intense electric fields (Fowler and Nordheim), 173.
Electrons in gases, experiments on motion of (Atkinson), 335.
Ellis (C. D.) and Aston (G. H.) The Dependance of the Photographic Action of β-Rays on their Velocity, 645.
Erosion by water-hammer (Cook), 481.
Eymers (J. G.) *See* Ornstein and others.

Films, thin, structure of, XII (Adam), 628.
Flow of water in rectangular pipes (Davies and White), 92.
Fluorescence of mercury vapour (Wood and Voss), 698.
Fowler (R. H.) and Nordheim (L.) Electron Emission in Intense Electric Fields, 173.
Frictional electricity, quantitative measurements (Macky), 107.

Gaseous combustion at high pressure, X (Newitt), 464.
Germanium, spark spectra (Rao and Narayan), 607.
Gibson (C. H.) and Hinshelwood (C. N.) The Homogeneous Reaction between Hydrogen and Oxygen, 591.
Greenhill (A. G.) Obituary Notice, i.
von Groth (P. H. R.) Obituary Notice, xx.

Hinshelwood (C. N.) *See* Gibson and Hinshelwood.
Hollingworth (J.) The Polarisation of Radio Waves, 444.
Homogeneous reaction between hydrogen and oxygen (Gibson and Hinshelwood), 591.
Hydrogen, Stark effect of fine structure (Schlapp), 313.

Intensity measurements in the secondary spectrum of hydrogen (Ornstein and others), 83.
Intermetallic compounds of alkali metals, spectra (Walter and Barrett), 257.
Ironside (R.) Diffraction of Cathode Rays by Thin Films of Copper, Silver and Tin, 668.

Jackson (W. W.) *See* Taylor and Jackson.

Kapitza (P.) The Study of the Specific Resistance of Bismuth Crystals and its Change in Strong Magnetic Fields and some Allied Problems, I, 358 ; II, 387 ; III, 401.
Kapucinski (W.) *See* Ornstein and others.
Kinetic method in the new statistics (Nordheim), 689.

Light of the night sky, III (Rayleigh), 11.
Lowry (T. M.) and Owen (G.) The Mechanism of Chemical Change, I, 505, 710.
Lowry (T. M.) and Vernon (M. A.) An Improved Method of Ultra-Violet Polarimetry—Anomalous Rotary Dispersion of Sodium Tartrate, 706.

Macdonald (H. M.) Note on Total Reflexion of Electric Waves at the Interface between Two Media, 523.
Macky (W. A.) On Quantitative Measurements in Frictional Electricity, 107.
Magnetic fields, strong, and specific resistance of bismuth crystals (Kapitza), 358, 387, 401.
Magnetic variation, diurnal, line-integrals (Price and Chapman), 182.

Mechanism of chemical change, I (Lowry and Owen), 505, 710.

Menzies (A. C.) The Spark Spectrum of Copper, 249.

Mercury, band spectra (Rayleigh), 349.

Mercury vapour fluorescence (Wood and Voss), 698.

Mittag-Leffler (G.) Obituary Notice, v.

Molecular streams, condensation on surfaces (Cockcroft), 293.

Mutarotation of tetra-acetylglucose (Lowry and Owen), 505.

Narayan (A. L.) *See* Rao and Narayan.

Neon, temperature and viscosity (Edwards), 578.

Newbery (E.) Metal Overvoltage Measurements with the Cathode Ray Oscillograph, 680. A Revision of the Theory of Transfer Resistance, 686.

Newitt (D. M.) Gaseous Combustion at High Pressures, X, 464.

Night sky, light, III (Rayleigh), 11.

Nordheim (L. W.) On the Kinetic Method in the New Statistics and its Application in the Electron Theory of Conductivity, 689.

Nordheim (L. W.) *See* also Fowler and Nordheim.

Obituary Notices :—
- Arrhenius (S. A.), ix.
- Dines (W. H.), xxiii.
- Greenhill (A. G.), i.
- von Groth (P. H. R.), xx.
- Mittag-Leffler (G.), v.

Ornstein (L. S.), Kapucinski (W.) and Eymers (J. G.) Intensity Measurements in the Secondary Spectrum of Hydrogen, 83.

Overvoltage, measurements by oscillograph (Newbery), 680.

Owen (G.) *See* Lowry and Owen.

Paget (Sir Richard) The Origin of Speech, 157.

Photo-Electric properties of metals in the soft X-ray region (Davies), 543.

Photometer for Compton Effect (Raman and Sogani), 526.

Polarimetry, ultra-violet (Lowry and Vernon), 706.

Powell (C. F.) Condensation Phenomena at Different Temperatures, 553.

Price (A. T.) and Chapman (S.) On Line-Integrals of the Diurnal Magnetic Variation, 182.

Radio waves, polarisation (Hollingworth), 444.

Raman (C. V.) and Sogani (C. M.) A Critical-Absorption Photometer for the Study of the Compton Effect, 526.

Rao (K. R.) and Narayan (A. L.) On the Series in the Spark Spectra of Germanium, 607.

Rayleigh (Lord) The Light of the Night Sky, III, 11. Observations on the Band Spectra of Mercury, 349.

Rayleigh's principle applied to continuous systems (Temple), 276.

Reflexion of electric waves at interface between two media (Macdonald), 523.

Reid (A.) Diffraction of Cathode Rays by Thin Celluloid Films, 663.

Richardson (O. W.) The Emission of Secondary Electrons and the Excitation of Soft X-Rays, 531.

Richardson (O. W.) and Chalklin (F. C.) The Excitation of Soft X-Rays, II, 60.

Saha theory and alkali metal vapours (Wilson), 1.

Schlapp (R.) The Stark Effect of the Fine Structure of Hydrogen, 313.

Schonland (B. F. J.) The Scattering of Cathode Rays, 673.

Sodium tartrate, anomalous rotary dispersion (Lowry and Vernon), 706.
Sogani (C. M.) *See* Raman and Sogani.
Spectra, band, mercury (Rayleigh), 349.
Spectra of alkali metals and their alloys (Walter and Barrett), 257.
Spectrum, secondary, of hydrogen, intensity measurements (Ornstein and others), 83.
Spectra, spark, of germanium (Rao and Narayan), 607.
Spectrum, spark, of copper (Menzies), 249.
Speech, origin (Sir Richard Paget), 157.
Stark effect of fine structure of hydrogen (Schlapp), 313.

Taylor (W. H.) and Jackson (W. W.) The Structure of Cyanite, 132.
Temple (G.) The Theory of Rayleigh's Principle as applied to Continuous Systems, 276.
Thin films, structure, XI (Adam), 628.
Thomson (G. P.) Experiments on the Diffraction of Cathode Rays, II, 651.
Timms (G.) The Nodal Cubic Surfaces and the Surfaces from which they are derived by Projection, 213.
Transfer resistance, revision of theory (Newbery), 686.

Ultra-violet polarimetry (Lowry and Vernon), 706.

Vernon (M. A.) *See* Lowry and Vernon.
Viscosity of neon, effect of temperature (Edwards), 578.
Vortices and resistance solids to moving in liquid (Bickley), 146.
Voss (V.) *See* Wood and Voss.

Walter (J. M.) and Barrett (S.) The Existence of Intermetallic Compounds in the Vapour State. The Spectra of the Alkali Metals and of their Alloys with each other, 257.
Water-hammer erosion (Cook), 481.
White (C. M.) *See* Davies and White.
Wilson (H. A.) The Saha Theory and the Conductivity of Flames containing Alkali Metal Vapours, 1.
Wood (R. W.) and Voss (V.) The Fluorescence of Mercury Vapour, 698.

X-ray region, soft, properties of metals (Davies), 543.
X-rays, soft, excitation (Richardson and Chalklin), 60.
X-rays, soft, excitation of and emission of secondary electrons (Richardson), 531.

END OF THE ONE HUNDRED AND NINETEENTH VOLUME (SERIES A).

I.A.R.I. 75

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY, NEW DELHI.

Date of Issue	Date of Issue	Date of Issue
25.5.61		
[illegible]		
24.1.64		
18.3.64		
5-1-66		
26 APR 1969		
- 6 FEB 1970		
7.10.77		
17 OCT 1977		

GIPNLK—H-40 I.A.R.I.—29-4-55—15,000